

Psychological Review

THEODORE M. NEWCOMB, Editor

UNIVERSITY OF MICHIGAN

Lorraine Bouthilet, Managing Editor

CONTENTS

Harvey A. Carr 1873-1954..... 81

Conditioning and Perception.....GREGORY RAZRAN 83

Punishment: II. An Interpretation of
Empirical Findings.....JAMES A. DINSMOOR 96

The Dynamics of Identification.....NEVITT SANFORD 106

A Statistical Model for the Process of
Visual Recognition.....ARNOLD BINDER 119

The Innsbruck Studies on Distorted
Visual Fields in Relation to an Or-
ganismic Theory of Perception.....HEINZ WERNER
AND SEYMOUR WAPNER 130

Since Learned Behavior Is Innate, and
Vice Versa, What Now?.....WILLIAM S. VERPLANCK 139

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

SOLOMON E. ASCH
ROBERT R. BLAKE
STUART W. COOK
CLYDE H. COOMBS
LEON FESTINGER
W. R. GARNER
JAMES J. GIBSON
D. O. HEBB
HARRY HELSON
E. R. HILGARD
CARL I. HOVLAND
E. LOWELL KELLY
DAVID KRECH
ROBERT W. LEEPER

ROBERT B. MACLEOD
DAVID C. MCCLELLAND
GEORGE A. MILLER
GARDNER MURPHY
OSCAR OESER
CARL PFAFFMANN
CARROLL C. PRATT
DAVID SHAKOW
RICHARD L. SOLOMON
ELIOT STELLAR
S. S. STEVENS
ERIC TRIST
EDWARD L. WALKER
ROBERT W. WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STREETS, LANCASTER, PA.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

\$6.50 volume

\$1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 3440, P. L. & E. of 1948, authorized Jan. 8, 1948

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 25th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Copyright 1955 by the American Psychological Association, Inc.





HARVEY A. CARR

THE PSYCHOLOGICAL REVIEW

HARVEY A. CARR

1873-1954

The death of Harvey A. Carr on June 27, 1954, at the age of eighty-one years has removed a familiar and influential figure from the psychological scene. Mr. Carr had been in retirement since 1938, when he became professor emeritus, and in this period of retirement he contributed relatively little to the field of psychology. Hence, in order to get some conception of his true professional stature, it will be necessary to focus on his activities during the 30 years (1908-1938) when he was on the faculty of the University of Chicago.

Mr. Carr, with BS and MS degrees from the University of Colorado, entered the University of Chicago as a graduate student in 1902. He came there with the intention of specializing in psychology, having had his interest in the discipline aroused by several years of teaching experience (he has asserted that, for instance, McMurray's text, *Methods of the Recitation*, profoundly influenced his thinking) and by his contact with Arthur Allin, who was his mentor at the University of Colorado. In his student days at Chicago Mr. Carr sat at the feet of James Rowland Angell and John B. Watson, each of whom became his close friend and left a deep impression on his thinking and values.

In 1905 Mr. Carr received his doctorate from the University of Chicago. Following this achievement he spent

about a year teaching in high school and two years as an instructor in psychology at Pratt Institute. It was at Pratt that he met the lovely and gracious woman who became his wife and lifelong companion.

He returned to the University of Chicago in 1908 and there rose from the rank of assistant professor to that of professor and chairman of the Department of Psychology, which latter post he held from 1926 to 1938. He was honored by being elected to the presidency of the American Psychological Association in 1926 and to a similar office in the Midwestern Psychological Association in 1937. In addition to his service in these presidential roles, he functioned on many national and local professional committees, was for a time associate editor of the *Journal of Experimental Psychology* and of the *Journal of General Psychology*, and held the editorship of the psychological series in the publications of Longmans, Green and Company.

Mr. Carr dubbed himself an experimental psychologist, though hastening always to make clear that he believed the classification of scientist was not restricted to experimentalists. His writings are many (somewhat over 50 units) and, while varied in scope, are concentrated in the fields of comparative psychology, space perception, and learning. Among his works are two books—

a once widely used introductory text on general psychology, *Psychology, a Study of Mental Activity* (1925), and *An Introduction to Visual Space Perception* (1935).

The publication order of his writings reflects well his maturing as a psychologist. His earlier publications are chiefly reports of experiments, while his later works are concerned more with basic concepts, e.g., the concept of learning, of the individual, of ability, of trait, and of reliability in measurement. His two books, published relatively late in his career, represent the integration of the products of his years of absorption with things psychological. A textbook on learning on which he labored never reached the publication stage because of the death of the co-author.

To a large percentage of the hundred and fifty or so students who took their doctorates in psychology during Mr. Carr's association with the Department of Psychology at the University of Chicago he served as advisor; about half of these students worked out their theses under his supervision.

His psychology was a brand of functionalism, much colored not only by the formulations of J. R. Angell and Wm. James, but also of G. F. Stout, Stanley Hall, J. B. Watson, and G. H. Mead. Mr. Carr was the last of the major textbook writers for about 25 years to include a chapter on the self in his introductory psychology. In making adaptation the focal concept of his system, in

lending sympathetic attention to the contributions from all fields of applied psychology, in emphasizing the correlations of molar behavior and physiological processes, and in giving extended treatment to telic processes, his functionalism adumbrated the form of present-day psychology. Also in the place which he gave learning in his system of psychology he was ahead of his time. Because of the breadth of his interests and his tolerance for new methods and points of view, he played a significant role in facilitating the many-sided development of psychology.

Mr. Carr was a Lincolnian figure in appearance and attitudes—lean, no dandy, humble, forthright, courageous, keen, judicious, generous in his efforts in behalf of others, and possessed of a good sense of humor as well as of warm sympathies. The welfare of his students was for him a major concern; his contacts with them, a source of great satisfaction. He was readily accessible to students, knew them as individuals, often published jointly with them, and treated them as equals. His disciples, as one might expect in the case of such a man, are many. Significantly, his disciples cannot be recognized by the specific content of their preachments or by the areas of their activity, but rather by the attitude with which they approach psychological problems. No higher tribute can be paid a teacher.

HELEN L. KOCH

University of Chicago

CONDITIONING AND PERCEPTION

GREGORY RAZRAN

Queens College

Woodworth's thoughtful article on "Reinforcement of Perception" (62) has for some time tempted the writer to join in a discussion of a timely topic. The temptation was "reinforced" recently when the writer went through a large volume of reports of Russian post-Pavlovian experiments on non-salivary visceral, and on what the Russians call "interoceptive," conditioning. The reports of the experiments, if one ignores customary introductory and concluding political dedications, are clearly pertinent to conditioning and perception and, while somewhat striking, are by no means wholly, if at all, at variance with what might be expected. Moreover, a good number of non-Russian experiments in this area have of late—it seems to the writer—been grounded by a perception-preoccupied academic psychology, and are in need of fresh integrations. And lastly, almost all of the writer's own unpublished CR experimentation bears directly on the relation of conditioning to perception.

However, before discussing the main thesis of conditioning and perception as raised by Woodworth's article, it seems relevant to take up a side issue, namely, Woodworth's classing of Pavlov as an advocate of a reinforcement view of conditioning.

As far as the writer is aware, Pavlov in no way ever considered reinforcement a means or a mechanism of the formation of conditioned reflexes. In the *Lectures on Conditioned Reflexes* (36), the developmental story of Pavlov's work and views, the word *reinforcement* occurs only once and then only as an incorrect translation of the

Russian.¹ In *Conditioned Reflexes* (35) reinforcement does occur and occurs often, mostly in the body of tables as "reflex reinforced" or "stimulus reinforced." But then, as the etymology of the word suggests, it has reference only to a means of strengthening (by repetition) an already formed conditioned reflex (or an *already conditioned stimulus*), and not to a means of setting up one. The fact is that Pavlov and his students have never even used the term reinforcement for trials *preceding* the full-fledged emergence of the CR, although the writer must plead guilty that he did apply the term—in pre-Hullian days, in 1929 (50), with no precognition of its ultimate import (grafting Thorndike upon Pavlov)—to all CR training trials. But he, too, was thinking then—and now—of nothing but the strengthening by repetition of an already formed hypothetical subliminal connection. And he also considered the convenience of saying "A CR was established after 20 reinforcements" rather than "A CR was established after 20 repeated combined applications of the conditioned and the unconditioned stimuli."

¹ The Russian sentence "Aeto vremennoye otosheniye i yego pravilo—usilivatsya s povtoreniyem . . ." (38, Vol. 3(1), p. 34) is translated as "This temporal relation and its law (reinforcement by repetition . . .)" (36, p. 56). It should be "This temporal relation and its rule—to become stronger by repetition. . . ." The seventh word in the Russian sentence, *usilivatsya*, by no means signifies "reinforcement," the Russian equivalent of which is *podkrepleniye* or *podkreplyatsya*. Compare the correct German translation: "Diese temporären Beziehungen und ihre Regel—das sie sich bei steter Wiederholungen verstärken . . ." (34, p. 18).

Pavlov's position on the mode of CR formation is really—despite some significant divergencies—much like that of Guthrie (in a somewhat modified Spence taxonomy, it should be classed as a contiguity neural S-S type). The terms that Pavlov uses a great deal in accounts of the genesis of the CR are "linkage" and "linking up"—Russian *zamykanye* and *zamykat*; Guthrie, it will be remembered, says that conditioning is like "setting a switch" (16, p. 81). Indeed, the Russian *zamykat* which, in addition to "linking up," also means "locking up," "hasping," "closing a circuit," and "setting a stone upon a vault," is about perfect for Guthrie's analogy. The difference between Guthrie and Pavlov comes, first, in what happens after the switch is set. Guthrie maintains that conditioning is not like "the wearing of a path" (16), implying that once set, the switch, unless interfered with, will stay set. But Pavlov, like a Russian, fears reversible reactions and insists that the set switch must be repeatedly tightened—and again it is to the tightening and not to the setting that reinforcement is applied. Pavlov also resorts on two occasions to the German term *Bahnung*, which, however, to him means not facilitation, but in his own words, and closer to its etymological meaning, "the laying down of fresh physiological paths in the centres [brain]" (35, p. 26)—and fresh paths must obviously be trodden to be kept in condition.

The last sentence presages another difference between Guthrie and Pavlov, namely, what it is that gets linked or switched. Pavlov, like Hebb (19), upholds neural cell assemblies ("analyzers"), whereas Guthrie adheres to movements and "movement-produced stimuli" (presumably, too, the neural correlates of the latter). Again, as is known, Pavlov's neural theories, postulating among other things that "strong

neural centers attract weaker ones," contain a good deal of James' and McDougall's "drainage."

And here a sentence in Pavlov's article in *Psychologies of 1930*—often repeated in his other writings—might be quoted to silhouette his main position (the writer uses the recently published Russian original of the article, and his translation differs a little from that in the *Psychologies*). The sentence reads: "... one must admit as a general phenomenon that in the higher region of the central nervous system every strongly excited center somehow directs to itself every other weaker excitation that happens to be present in the system at the same time, and in this manner the point of the application of the [weaker] excitation and the center become more or less firmly connected for a definite time and under definite conditions (law of neural linkage—association)" (37, p. 210). The words "law of neural linkage—association" (Pavlov uses in Russian *assotziatziya* with its specific connotation) seem to the writer particularly significant, since in personal conversation Pavlov definitely expressed himself as "an objective and experimental" continuer of Hartleyan associationism, mentioning even that he, too, was an M.D.; and in his very last article (1935) he plainly stated: "Conditioned connections, as shown earlier, obviously are what is known as association through simultaneity [contiguity]. Generalizations of conditioned connections correspond to what is known as association through similarity. The synthesis and the analysis of conditioned reflexes (associations)—are essentially the same basic processes of our mental work [life]" (38, Vol. 3[2], p. 345). "Rewards" and "needs" were essentially foreign to Pavlov's thinking, as they were to Hartley's. Pavlov, like Freud, knew very little about modern behavioral psychology, and he conceived

of the dynamics of neurobehavior not as a system in which strong centers (or reactions) reinforce and aid the weak centers to become independent, but as one in which the strong centers direct the weak ones to help in expanding the realm of the strong.

But let us take up now Woodworth's main thesis. This states, as will be remembered, that "The conditioning experiment is really concerned with the establishment of a new perception . . ." (62, p. 124) and "The new learning, the conditioning, is sensory and not motor . . ." (p. 122). (Strictly speaking, Woodworth means of course that the conditioning is *postsensory*, since perception to him is—and always has been—" . . . an elaboration of the receptive process, adjusting the organism to the objective situation rather than to the stimuli received . . ." [p. 122].) Perception—as well as expectancy—is furthermore physicalized into "registration," and it is the registration that becomes reinforced in a "trial-and-check process" of "two phases": a trial phase of "tentative characterization" and a check phase that is " . . . an acceptance or rejection, a positive or negative reinforcement of the tentative perception . . ." (p. 124). Moreover, " . . . perception is always driven by a direct, inherent motive which might be called the will to perceive . . ." (p. 123).

If labeling is desired—and to the writer labeling is more a matter of taxonomy and pedagogy than of ontology—then Woodworth seemingly suggests a "central S-S reinforcement theory." (Reinforcement of perception by the action of the *peripheral* portion of the unconditioned agent is presumably not implied since conditioning is well effected without such peripheral participation [12].) As such, the theory's thesis—or synthesis—not only contains a good deal that is common to the

two main wings of contemporary learning theorists—Pavlov, Lashley, Tolman, Guthrie, and Hebb, on the one hand and Thorndike, Hull, and Skinner, on the other—but is also in line, at least not out of line, with a considerable number of current conditioning studies. Central S-S theories have long been favored by neurophysiological studies on loci of conditioned modifications (18) (also supported by White and Schlosberg's recent experiment [61]), and central reinforcement views are among other things not open to the objections of Brogden's sensory preconditioning (7) and Sheffield's "no need reduction" learning (54). Moreover, perceptions that are driven by their own motives obviously reduce the distance between rewards and emphases and indeed also between reinforcement and contiguity (they may even make the concept of reinforcement altogether superfluous), while trials and checks are good mediators between trials and errors and hypotheses. In short, the theory covers a lot of psychological territory, and an analysis of its experimental base would seem desirable, and timely.

Generally speaking, such an experimental analysis should properly begin with a delimiting definition of the two terms involved: conditioning and perception. But here the definitions of the terms differ so widely—particularly that of perception²—that it might be best

² The divergence in the delimitation of perception might perhaps be illustrated by a simple discrimination test. If a group of perception theorists were asked to divide into perceptual and nonperceptual the entire hierarchy of human reactions—ranging, let us say, from the simplest viscerovisceral reflexes to reactions guiding the writing of a scientific treatise—it is likely that the area of disagreement in the middle of the hierarchy would exceed the area of agreement at the extremes. Offhand, the writer can think of such a divergent delimitation as one that would limit perception to the "awareness or belief of the truth of a proposition" (No. 4 in Warren's

to confine the discussion, at least at first, to clear-cut cases of each. Consequently, conditioning will refer only to the so-called classical variety in which a definite unconditioned (typically dominant or prepotent) motor and/or glandular reaction (typically a set of reactions) and its stimulus are paired one or more times with a definite to-be-conditioned stimulus and reaction, and in which the resulting modifications can be attributed to nothing else but the "mere" pairing of the to-be-conditioned and unconditioned stimuli and reactions. Similarly, perception and perceptual reactions will relate only to highly conscious and plainly adjustive registrations and judgments and *Kundgabe* which psychologists who use the term would hardly hesitate to designate as perception or perceptual; without perception and nonperceptual will denote only those reactions for which the assumption of perception would admittedly be gratuitous, e.g., leucocytosis, bile secretion, splenetic contraction, kidney action, extrasystole, and the like—as unconditioned reactions; subliminal interoceptive stimulation as to-be-conditioned stimuli; and reactions in spinal animals, in animals very low in the phyletic scale, and in somnambulistic hypnoses. (When, however, for human subjects, reference is made not to perceptions of stimuli and reactions by themselves, but to *perceptions of relations* between the stimuli and reactions, then perception will mean specific knowledge of the relations—as determined by instructions and verbal reports—and no perception will signify total absence of awareness of the relations.) The restriction of the two terms is for the purpose of illustrative clarity and logical consistency and is by no means intended to bar wider concep-

tionary), as distinct from one that would extend it to all reactions to exteroceptive stimuli.

tions of either. Indeed, it is hoped that what will be said here about the relation of conditioning to perception will not differ much from what might be said about the relation of learning to perception, no matter how each is defined.

Thus phrased, the analysis of the relation of conditioning to perception might well proceed by dividing the entire problem into four specific questions:

1. Do nonperceptual reactions become conditioned? Does conditioning occur when either the unconditioned reaction or the to-be-conditioned reaction (the original reaction to the to-be-conditioned stimulus) or both are nonperceptual? This is the question of *conditioning without perception*.

2. Do perceptual reactions become conditioned? Does conditioning occur when the unconditioned and the to-be-conditioned reactions are both perceptual but the relation between them is nonperceptual? And, if such conditioning takes place, in what ways, if any, does it differ from the conditioning of nonperceptual reactions? This is the question of *conditioning of perception*.

3. What is the course of conditioning when the relation between the two reactions, in addition to the perception of each reaction individually, is clearly perceived? Or, what is the role of perception in conditioning? This is the question of *conditioning with perception* or *conditioning through perception*.

4. May mere conditioning give rise to new—and novel—perceptions? Transform peripheral and nonadjustive into central and adjustive reactions? Or change receptive into postreceptive registrations (Woodworth)? Or convert presensory into sensory reflexes (Bykov)? Or permit, in general, new reorganizations, insights, or *Gestaltungen* in already perceived reactions? This is the question of *perception through conditioning*. (A fifth question of *perception without conditioning* is beyond the scope of the present article.)

Attempted answers will be presented in the same order as the questions:

1. *Conditioning without perception*.

As already indicated, Russian laboratories have in recent years reported a large number of successful experiments on the conditioning of purely vegetative functions. And while the experiments need verification, the high consistency of

their results leaves, other things being equal, little room for doubting their main findings. The fact is that, judging from published reports, Russian post-Pavlovian conditioning studies seem to be better designed, more carefully controlled, and more ingeniously recorded and measured (even using statistics) than those that came out of Pavlov's laboratory during his lifetime—and the recent "canonization" of anything that Pavlov said bulwarks well against the accretion of new neural speculations. The studies pertinent to the present problem are primarily of two types. First are a few dozen separate experiments on what might be called exteroceptive-visceral conditioning; typical Pavlovian bells, buzzers, metronomes, whirligigs, light flashes, smells, touches, and the like have been conditioned to produce significant changes in a variety of reactions. Varieties of reactions included are (space permits only a sample citation of Russian experiments): cardiovascular (1, 33), respiratory (63), gastrointestinal (22, 57), glycometabolic (28, 51), hepatic (2), renal (3, 9), splenic (10), and even thermoregulatory (39), oxygenational (9, 60), leucocytotic (9) anaphylactic (9, 55), basal metabolic, (9) and those dealing with vitamin absorption (8). Secondly, almost as many studies can be found on the successful conditioning of viscerosomatic and viscerovisceral reflexes—or what the Russians call interoceptive conditioning—in which a pure internal visceral stimulation is conditioned to elicit some external or some other internal motor and/or glandular reaction (4, 9, 32). Both series of conditionings have been effected in both animal and human subjects, and to attribute perception in either case can hardly be warranted unless one prefers to equate perception with reaction, which, of course, destroys the very meaning of the category.

Again, the Russians have recently reported about half a dozen experiments on subliminal EDR conditioning (13), several successful studies on conditioning in somnambulistic hypnosis (25), and at least one clear-cut conditioning of pupillary contraction (14). (Conditioned contraction of the pupil, after a great many trials, is also claimed by a recent Japanese experimenter [24].) Then, subcortical conditioning has, of course, been readily obtained in both American and Russian laboratories for some time; and while the Russians continue to disclaim spinal conditioning, the writer is by no means convinced that their own results do not support it (53). Spinal conditioning is looked upon with disfavor in contemporary Russia as anti-Pavlovian: a recent lengthy article (11) in Russia's outstanding physiological periodical—*Fiziologicheskii Zhurnal SSSR*—was wholly devoted to taking sharp issue with the "reactionary-mechanistic-idealistic doctrines" of Shurrager and Culler (56) and to praising Kellogg *et al.* for their "proper unmasking" of Shurrager and Culler as followers of the "pseudo-scientific mystical-metaphysical Lashleyan views" (*sic!*). (Kellogg *et al.*: "Results like those of Culler and Shurrager follow as a logical sequence from the original experiments of Lashley" [21, p. 100].) Finally, it should be mentioned that some twenty years ago the writer summarized (41) unmistakable instances of conditioning in animals very low in the phyletic scale (Protozoa, Crustacea); since then a good many other experiments have fully corroborated and extended the earlier findings (Paramecia, 58; Coelenterata, 64; silkworms, 27). But for some reason psychologists have in recent years wholly neglected their ancestors in the Early Paleozoic Age.

The main generalization of the presented evidence—which, unfortunately, could be no more than mentioned here

—is obvious and hard to avoid. It is that in both animals and man there must exist at least a kind of modification through learning that is nonperceptual, that is—we might say—segmental, and even cellular, rather than central, and protoplasmic rather than perceptual. No other interpretation seems adequate as one goes through the vast literature of clear-cut modifications through conditioning of the very lowest and very simplest reactions that man and animal are capable of. Moreover, it would seem logical to assume that this nonperceptual modification—or conditioning—must in a large way be prototypic of all learning modifications, and that the question whether a particular learning is or is not nonperceptual should hinge to a large extent upon how close functionally the particular learning is to nonperceptual conditioning. And here Pavlov's intensive and extensive experiments on the salivary conditioning of dogs, the protocols of which have been reported in the fullest details and are known to the writer, would appear to be the most suitable area to study and compare.

Yet all that the distinctive features of salivary canine conditioning—as revealed in the protocols—can possibly warrant is the postulation of an extra factor, that might be called perception, to account for the extinction and the generalization of this conditioning. Both extinction, which is very fast in salivary conditioning, and generalization, which is really discrimination, are as a rule difficult and often altogether impossible in interoceptive and subcortical and “low phyletic” conditioning. And neither phenomenon has been satisfactorily conceptualized (both the “extinctive inhibition” of Pavlov and the “reactive inhibition” of Mowrer-Miller-Hull are more like a Hegelian dialectic than an empirical systematization). Elsewhere (48) the writer sug-

gested a factor of “perceptual categorizing” in generalization, and here he would add that extinction, too, may involve a factor of “perceptual check” or “relational learning” (colloquially, an “aha” reaction when the conditioned stimulus is not what it used to be). Grant's recent study (15) lends some support to such a view, the mentioned Russian studies on interoceptive-exteroceptive CR interactions offer a good basis for it, and there is an old but significant finding by Boldyreff (6) that deoesophagized dogs do not extinguish their salivary CR's.³ But all this does not pertain to the *mere acquisition* of conditioned reactions for which, the writer repeats, perception is unnecessary and for which, if anything, it may be more an obstacle than a vehicle.

2. *Conditioning of perception.* In the last fifteen years, the writer's typical experimental design in conditioning has been to misinform his subjects about the true nature of the experiment in which they participated. These subjects were fully cognizant of both the unconditioned and the to-be-conditioned stimuli reactions but were prevented by the instructions from perceiving the pertinent relations—the conditioning possibilities—between the two. The unconditioned act was in all experiments the consumption of food but the to-be-conditioned stimuli and reactions varied widely: stimuli varied from metronomes to musical selections, from light flashes to paintings, and from nonsense syllables to literary quotations; reactions included salivations and hunger sensa-

³ It is by no means contended that “there is no extinction without perception,” but merely that extinction—as well as generalization—readily permits perceptual factors to operate and that they do operate whenever perception is available. Other factors in extinction are not denied and, except for emphasis, the writer still adheres to the view which he expressed previously in this journal (47).

tions, affective ratings and adjectival characterizations, and speeds of unscrambling food-related words and frequencies of finding food-related rhymes. But in all cases, modifications in the subjects' perceived but perceptually unrelated reactions were clearly formed and formed in a manner not unlike that of nonperceptual conditioning: arbitrary, uncognized, and—most probably—unadjustively “registered.” Most highly cognized reactions such as complex characterizations of most meaningful music and “rational” ratings of most “ego-involved” political slogans could well be modified merely through food consumption in a simple conditioning fashion provided the relation between modifier and modified was unperceived. The crux of the course of the conditioning—whether it was to be simple conditioning or something else—was at all times *not the perception of the individual reactions and stimuli in the conditioning situations, but the perception of the relation between them.*

To be sure, differences between the writer's conditionings of perceptions and the Russians' conditionings of vegetative functions did exist and could be noted. For one thing, the writer was continually confronted with the problem of selectivity in the conditionings, of “just what gets conditioned.” More exactly, it was a matter of conditioning priority manifesting itself in such findings that coherent sentences, semantic aspects of words and sentences, simple tone ratios, and “well configured” patterns of light flashes were more readily and more fixedly conditioned than incoherent sentences, phonetic aspects of words and sentences, complex tone ratios, and “poorly configured” light patterns (44, 46, 49). Lashley's statement that “In any trial of a training series, only those components of the stimulating situation that are dominant in the organization [of any nervous activity or reaction]

are associated” (26, p. 242) was well borne out. The exceptions to Lashley's statement, if they are exceptions, were only that the nondominant components of the conditioning situations were, as a rule, associated “much less” rather than “not at all,” and that dominance was often determined much more by the subjects' attitudes and “activities in progress” (Woodworth) than by mere “compellingness of stimuli.” When the to-be-conditioned stimuli were flashes of miniature red and green lights which at the same time served as right and wrong cues in solving a bolt maze, the course of the conditioning was plainly more a function of the rightness and wrongness of the lights than of their spatial arrangements (43). For another thing, in the writer's own experiments, unmistakable differences between the conditionings of salivary and judgmental—or reflexive and attitudinal—reactions were plainly evidenced. Salivary conditioning was, for instance, helped but little by subjects' postconditioning verbalization of the CR situation, but in the conditioned modifications of judgments such verbalizations were very significant. Indeed, it is doubtful whether arbitrarily conditioned judgments could stay conditioned at all without some subsequent verbal consolidation or interiorization (sensory and “unconscious-attitude” conditioning do not seem to be subject so much to these restrictions). Yet, these judgments are conditionable in the laboratory—and probably are an everyday event in life situations—and the earlier statement that the crux of the conditioning-perception interaction inheres, not in the perception of the conditioned events, but in the perception of the conditioning relation between the events is essentially valid. The matter will be further considered in the coming section.

3. *Conditioning with perception.* Conditioning-like experimental arrange-

ments with perceived to-be-conditioned and unconditioned stimuli and reactions that were, or came to be, also perceived as related to each other, form perhaps the bulk of the experimental literature on human conditioning. In most of the experiments, the subjects themselves "caught on," after one or several trials, to the pertinent relationship; but in some cases, as in the studies by Hilgard and by the writer, the subjects were either just told the relationship or asked to facilitate or inhibit it. No attempt at all will be made here to cite and summarize even most briefly the results of these experiments, except to say that they certainly are not those of mere conditioning and that only a highly sophisticated psychologist could fail to notice that. The fact is that most of these "conditioning with perception" experiments should not even be classed as studies in learning, both because their results have been influenced so much more by such factors as suggestion, alertness, concern, and the like than by practice itself, and because of their extreme irregularity and near-zero efficiency; some subjects became conditioned in one or two trials, others took hundreds or even thousands of trials, and still others were totally unaffected by the conditioning. More analytically, the reported experiments deviated from typical learning tasks in that they seem to have involved primarily, not the acquisition of new reactions, but the confrontation of old and new reactions—systems of old symbolic attitudes vs. arrays of new somatic stimuli—so that their unintended results are more significant than their intended ones, and their conditioning against perception more revealing than their conditioning with perception. At any rate, it would probably not be an overstatement to say, for instance, that the more than a hundred experiments on the development of conditioned faradic withdrawal

responses, particularly the protracted and laborious ones that came out of Russian laboratories, have by themselves contributed very little indeed to either learning theory or learning practice (42).

Again, conditioning in which the pertinent relations between the stimuli and the reactions are perceived is really discrimination and relational learning rather than mere conditioning. Fifteen years ago in an article in this journal on "The Transposition of Relational Responses and the Generalization of Conditioned Responses" (45), the writer expressed the view that relational learning is a separate *sui generis* category of modifiability the characteristics of which cannot be deduced from those of mere conditioning; a number of well-controlled American experiments performed since then (5, 17, 31) would seem to reaffirm this view (itself grounded in the work of Lashley and Beritoff). The view also implied that relational learning is a higher and more complex form of learning than mere conditioning and thus would be expected to be used by organisms capable of using it. Lloyd Morgan's principle that higher forms of activities should not be attributed to a task when lower forms could accomplish it, is really a law of parsimony for experimenters and theorists and not for subjects and performances. Higher forms are also more efficient forms and will be resorted to when needed. In short, it is the writer's contention that conditioning with perceived relationships is neither "mere conditioning" or "conditioning plus" but something else: it is relational or perceptual learning.

4. *Perception through conditioning.* The two clinical instances, one reported by Hilgard and Marquis (20) and the other by Sears and Cohen (52), of subjects who regained cutaneous sensitivity in their hysterically anesthetized left hands by means of a CR technique, are

no doubt too well known to require detailing. The cases were, of course, only clinical and had no controls. However, the Russians seem to have embarked recently on a whole series of experiments on what they call "transformation of presensory into sensory reflexes," and two reports of such experiments, both by Pshonick, are so far available. In one of the experiments (9) Pshonick showed that when a nocuous but presensory (subliminal) chemical stimulation of the wrist that produced vasodilatation is repeatedly paired with the flashing of a blue light, the resulting CR to the light is at first also presensory and vasodilative, but in the course of training it becomes sensory-arousing definite algesis—and vasoconstrictive. Yet, the unconditioned chemical stimulation itself continues to elicit the original reaction. In the other experiment (40), Pshonick proved that conditioned thermalgesias and thermesthesias not only persist after the unconditioned points of stimulation are anesthetized, but that, in general, conditioned sensations may come to attain dominance over unconditioned ones. Thus, when an unconditioned thermal stimulus (43°C.) was followed after ten seconds by an unconditioned thermalgesic one (63°C.), the sensory result was, as might be expected, a sensation of heat and pain succeeding one of warmth, and there was also a change from vasodilatation to vasoconstriction. But when a conditioned thermal stimulus such as the ringing of a bell was first applied, the sensation of warmth and the vasodilatation persisted in the face of the subsequent activation of the unconditioned thermalgesic stimulus (63°C.), but was readily stopped and transformed into a sensation of heat and vasoconstriction by a subsequent conditioned thermalgesic stimulus such as the flashing of a light.

In a different area, the writer has

shown that very complex but specific characterizations of a variety of musical selections and paintings could attain very much "higher levels of perception" by merely presenting the selections and the paintings one or more times while the subjects are enjoying a free lunch. The characterizations were expressed by means of a standardized check list of 80 adjectives and by extensive freely written statements; and an analysis of both clearly indicates that the conditioned resultant of the interaction of the affectivity of the food with the affectivity of the music and the paintings is almost always a higher perceptual integration, really an emergent. The association with—or the conditioning by—the food and the eating (subjects were of course never aware of the nature of the experiment) unmistakably increased, as determined by clear-cut objective indices, the subjects' objectivity of the characterizations, agreement with expert opinion, and stimulus-boundness or intrinsicity (that is, characterizations uninfluenced by labeling and stereotypy). The writer named this form of modification *emergent conditioning* but it really is *emergence through conditioning*. Emergence and novelty have been, as is known, emphasized especially by one group of psychologists, namely the Gestaltists, who are largely not too mindful of "pasts" and antecedents. But it is difficult to see how anything can emerge from nothing, and why novelty and association—or conditioning—cannot be complementary rather than excluding concepts, or why the meaningful should not arise from the meaningless (it must arise from somewhere). A world where all is "given" meaning appears, ontologically, meaningless and, pragmatically, too heavenly to be human. Indeed, the Lord Himself may well be irritated by Man's constantly and unremittingly perceiving and meaning and

problem solving and cognizing and creating. The Lord rested after He had created.

FINAL COMMENT

What the writer has suggested here—upon the basis of an array of experimental evidence—is obviously a two-level theory of learning. More exactly, it is *at least* a two-level theory since, on the one hand, the writer is not quite ready to equate perception with relational learning (nor willing to abandon his previous distinction between contemporaneous and biographical, or time-binding, learning); and, on the other hand, a tenable argument can be made—and has been made—for two distinct classes of conditioning. Nonetheless, at least for the time being, the theory may well be regarded as a dual one. It believes the widest *sui generis* division to obtain between what for the present is best described as a perceptual and a nonperceptual level of learning, and it asserts that the chief determinant of the division is the presence or absence of perceived—or reacted—relations between stimuli and reactions involved in learning rather than the mere nature and history of the stimuli and reactions per se. The theory furthermore holds that *outside* reinforcement is not the *sine qua non* of learning and would prefer to discard the term altogether. The concept of contiguity supplemented by one of dominance-selection (or prepotency-selection) would seem to suffice.

Space does not permit any elaboration of the dominance-selection concept except to refer to Murphy (30) and to state that dominance determinants are specific functions of tasks, organisms, and biographies. But a word must be said about the levels of learning and their interactions. To begin with, the writer's levels of learning differ from "kinds of learning" (Tolman, 59) and from "factors of learning" (Mowrer,

29) in assuming that one kind of learning, the higher, dominates the other kind, the lower, and that the other kind, the lower, subsists in the higher. Or, in other words, whenever learning takes place, it involves the operation not of one kind *or* the other kind, but of either (a) the lower kind or (b) the higher *and* the lower kinds. And specifically, it obviously means that relational learning subsumes conditioning, and biographical learning—if regarded as another true level—subsumes both relational learning and conditioning. Two examples might serve as illustrations: (a) Human subjects when they "catch on" to the S-R relations in a CR experiment greatly modify thereby their conditioning but do not as a rule wholly nullify it (some mere conditioning often occurs even if the subjects "decide" or are instructed to oppose attempts to condition them). (b) In the writer's series of experiments on configural conditioning—both with light flashes and with tonal triads—*perceived patterns* certainly determined largely the course of modifiability but *unperceived parts* were by no means completely left out. The rationale in both cases seemed to be that while the higher level of learning dominated the lower level when both were functioning, the conditions for the functioning of the higher level were more complex and more disruptable, so that the lower level did at times gain *quantitative* superiority.⁴

Two-level (or even three-and four-level) theories of learning need, in the

⁴ The writer would class Harlow's "learning sets" (17) as biographical learning and the recent experiment (31) by Nissen *et al.* as relational and conditioning. It is interesting to note that the Russians had found salivary conditioning in monkeys, as compared with that of dogs, quite difficult and irregular (25), resembling in a number of ways the salivary conditioning of uninstructed human subjects for which the writer postulated sets some time ago.

writer's opinion, no general apology and are, the writer believes, inoffensive to thinking grounded in areas other than learning experiments. The view that learning existing on this planet for hundreds of millions of years has continued in one quantitative continuum without evolving any new qualitatively distinct level seems very unlikely and discouraging. And the view that all learning changes are only changes in perception and that our vast nonperceptual equipment is either unalterably fixed or unalterably subservient to perceptual learning is forsooth more depictive of satans or saints than of true humans. Academic psychology has really, for some time, been suffering not only from a chronic and stubborn reductionism "from below" but also from one "from above." The obvious consideration that just as, by itself, a specific motor reaction is a restricted segment of the "Whole of Man," so a perceptual reaction is by itself both spatially and temporally a restricted segment of the "Whole of Nature," is usually overlooked. And the need for combining approaches "from below" and "from above" is patently ignored. Fortunately, however, the last few years seem to bear some signs of a change as one feels a palpable rise of the middle-of-the-road and hears from all sides voices calling for syntheses and *rapprochements*. And while the attempts at synthesis may well be, as yet, very tentative and inchoate, there is a will to learn.

REFERENCES

1. ARINCHIN, N. A., & KARMENOV, I. G. [Conditioned reflex changes of venous pressure and tonus.] *Fiziol. Zh. SSSR*, 1953, 39, 594-600.
2. BALAKIN, S. L. [The highest nervous activity and the bile secreting activity of the liver.] *Fiziol. Zh. SSSR*, 1940, 19, 503-510.
3. BALAKSHIN, V. L. [The mechanism of conditioned reflex kidney action.] *Trud. Leningr. Gosud. Univer.*, 1936, 17, 61-108.
4. BELOUS, A. A., & GREBENKOV, M. A. [Conditioned reflexes of carotid chemoreceptors.] *Fiziol. Zh. SSSR*, 1953, 39, 591-593.
5. BITTERMAN, M. E., & WODINSKY, J. Simultaneous and successive discrimination. *Psychol. Rev.*, 1953, 60, 371-376.
6. BOLDYREFF, W. N. *Two new laws of cerebral function*. Battle Creek, Mich.: Battle Creek Sanitarium, 1931.
7. BROGDEN, W. J. Sensory pre-conditioning with human subjects. *J. exp. Psychol.*, 1947, 37, 527-539.
8. BUNYATIN, G. K., KECHECK, Y. A., & MATINYAN, G. V. [The effects of unconditioned and conditioned algesic stimuli on some aspects of the exchange of ascorbic acid in living organisms.] *Fiziol. Zh. SSSR*, 1951, 37, 225-231.
9. BYKOV, K. M. [The cerebral cortex and the internal organs.] Moscow: Medgiz, 1947.
10. BYKOV, K. M., & GORSHEV, M. A. [The formation of conditioned splenic movements.] *Vestn. Khirur.*, 1932, 27, 46-50.
11. DANILOV, I. M. [Concerning one American attempt to revise Pavlov's teachings.] *Fiziol. Zh. SSSR*, 1952, 38, 368-375.
12. FINCH, G. Salivary conditioning in atrophinized dogs. *Amer. J. Physiol.*, 1938, 124, 136-141.
13. GERSHUNI, G. V. [Reflex reactions during the action of external stimuli upon man's sense organs and sensations.] *Fiziol. Zh. SSSR*, 1949, 35, 541-559.
14. GLAZER, V. D. [Conditioned reflex pupillary contraction.] *Fiziol. Zh. SSSR*, 1953, 39, 571-579.
15. GRANT, D. A., & SCHIPPER, L. M. The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed-ratio random reinforcements. *J. exp. Psychol.*, 1952, 43, 313-320.
16. GUTHRIE, E. R. *Psychology of learning*. (2nd Ed.) New York: Harper, 1952.
17. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, 56, 51-65.
18. HARLOW, H. F., & BRONNER, J. A. Acquisition of new responses during inactivation of the motor, premotor, and somesthetic cortex in the monkey. *J. gen. Psychol.*, 1942, 26, 299-313.
19. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.

20. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: D. Appleton-Century, 1940.
21. KELLOGG, W. N., DEESE, N. J., PRONKO, N. H., & FEINBERG, M. An attempt to condition the chronic spinal dog. *J. exp. Psychol.*, 1947, 37, 99-117.
22. KOGAN, B. A. Ueber den Einfluss des bedingten Nahrungsreizes auf die exkretorische Pankreasfunktion. *Zsch. klin. Med.*, 1931, 117, 203-209.
23. KOROTKIN, I. I., & SUSLOVA, M. M. [A study of the highest nervous activity in somnambulistic hypnosis.] *Fiziol. Zh. SSSR*, 1953, 39, 423-431.
24. KOTAKO, Y., & MIHAMA, H. Conditioning of pupillary reflex in man. *Jap. J. Psychol.*, 1952, 22, 77-78.
25. KRIAZHEV, V. YA. [Salivary conditioned reflexes in monkeys.] *Fiziol. Zh. SSSR*, 1941, 30, 490-495.
26. LASHLEY, K. S., & WADE, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, 53, 72-87.
27. LOBASHEV, M. E., & NIKITINA, I. A. [Temporal connections in silkworms.] *Dokl. Akad. Nauk*, 1951, 79, 1057-1059.
28. MALEV, I. A. [Conditioned reflex hypoglycemia and its clinical significance.] *Klin. Med.*, 1951, 39(9), 41-42.
29. MOWRER, O. H. Two-factor learning theory: summary and comment. *Psychol. Rev.*, 1951, 58, 350-367.
30. MURPHY, G. *Personality: a biosocial approach*. New York: Harper, 1947.
31. NISSEN, H. W., LEVINSON, B., & NICHOLS, J. W. Reinforcement and "hypothesis" in the discrimination behavior of chimpanzees. *J. exp. Psychol.*, 1953, 45, 334-344.
32. OKHNYANSKAYA, L. G. [The study of the conditioned respiratory-vasomotor reflexes.] *Fiziol. Zh. SSSR*, 1953, 39, 610-613.
33. PAMER, I. A. [Conditioned extrasystole in man.] *Fiziol. Zh. SSSR*, 1953, 39, 286-292.
34. PAVLOV, I. P. *Die höchste Nerventätigkeit (das Verhalten) von Tieren*. München: J. F. Bergman, 1926.
35. PAVLOV, I. P. *Conditioned reflexes*. London: Oxford Univer. Press, 1927.
36. PAVLOV, I. P. *Lectures on conditioned reflexes*. New York: International Universities Press, 1928.
37. PAVLOV, I. P. A brief outline of the higher nervous activity. In C. Murchison (Ed.), *Psychologies of 1930*. Worcester: Clark Univer. Press, 1930. Pp. 207-220.
38. PAVLOV, I. P. [Collected works.] (2nd Ed.) 5 Vols. Moscow: Adakemiya Nauk, 1951.
39. POPOV, T. V. [The effects of caffeine and bromine on a weak and on a well established conditioned polypnoe.] *Fiziol. Zh. SSSR*, 1948, 34, 550-554.
40. PSHONIK, A. T. [The role of the cerebral cortex in the formation of cutaneous algesis.] In K. M. Bykov (Ed.), [Problems of cortico-visceral pathology.] Moscow: Akad. Meditz. Nauk, 1949. Pp. 33-55.
41. RAZRAN, G. Conditioned responses in animals other than dogs. *Psychol. Bull.*, 1933, 30, 261-324.
42. RAZRAN, G. Conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects. *Psychol. Bull.*, 1934, 34, 111-143.
43. RAZRAN, G. Attitudinal control of human conditioning. *J. Psychol.*, 1936, 2, 327-337.
44. RAZRAN, G. Studies in configural conditioning. VII. Ratios and elements in salivary conditioning to various musical intervals. *Psychol. Rec.*, 1938, 2, 370-376.
45. RAZRAN, G. Transposition of relational responses and generalization of conditioned responses. *Psychol. Rev.*, 1938, 45, 532-538.
46. RAZRAN, G. Studies in configural conditioning. IV. Gestalt organization and configural conditioning. *J. Psychol.*, 1939, 7, 3-16.
47. RAZRAN, G. The nature of the extinctive process. *Psychol. Rev.*, 1939, 46, 264-297.
48. RAZRAN, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, 46, 337-365.
49. RAZRAN, G. Experimental semantics. *Trans. N. Y. Acad. Sci.*, 1952, 14, 171-177.
50. RAZRAN, G., & WARDEN, G. J. The sensory capacities of the dog as studied by the conditioned reflex method. *Psychol. Bull.*, 1929, 26, 202-222.
51. SAVCHENKO, V. A. [Conditioned reflex hypoglycemia, glycosuria, and hyperglycemia.] *Bull. eksp. Biol. Med.*, 1940, 9, 293-295.
52. SEARS, R. R., & COHEN, L. H. Hysterical anaesthesia, analgesia, and astereognosis. *Arch. Neurol. Psychiat. N. Y.*, 1939, 29, 260-271.
53. SHAMARIN, N. A., & NESMEYANOV, T. N. [Experimental transformation of reflexive reactions of the spinal cord.] *Fiziol. Zh. SSSR*, 1953, 39, 601-609.

54. SHEFFIELD, F. D., WULF, J. J., & BARKER, R. Reward value of copulation without sex drive reduction. *J. comp. physiol. Psychol.*, 1951, **44**, 3-8.
55. SHUR, E. I., & KADYKOV, B. I. [Changes in blood precipitation under the influence of conditioned reflexes.] *Bull. eksp. Biol. Med.*, 1940, **10**, 191-193.
56. SHURRAGER, P. S., & CULLER, E. A. Conditioning in the spinal dog. *J. exp. Psychol.*, 1940, **26**, 133-159.
57. SINELNIKOW, E. I. Ueber den Einfluss des Grosshirnrinde auf die motorische Funktion des Dünndarms. *Arch. ges. Physiol.*, 1926, **213**, 239-244.
58. TCHAKHOTINE, S. Réactions conditionées par micropuncture ultraviolette dans le comportement d'une cellule isolée (*Paramecium caudatum*). *Arch. Inst. Prophyl.*, 1938, **10**, 119-133.
59. TOLMAN, E. C. There is more than one kind of learning. *Psychol. Rev.*, 1949, **56**, 144-155.
60. VOYTOKOVICH, V. I. [Conditioned reflex regulation of blood oxygenation.] *Fiziol. Zh. SSSR*, 1952, **38**, 452-458.
61. WHITE, L. T., & SCHLOSBERG, H. Degree of conditioning as a function of the period of delay. *J. exp. Psychol.*, 1952, **43**, 357-362.
62. WOODWORTH, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, **60**, 119-124.
63. ZBOROVSKAYA, I. I., & DOLIN, A. O. [Conditioned toxic dyspnea.] *Fiziol. Zh. SSSR*, 1939, **27**, 13-21.
64. ZUBKOV, A. A., & POLIKARPOV, G. G. [Conditioned reflexes in coelenterates.] *Usp. sovr. Biol.*, 1951, **32**(5), 301-302.

(Received March 19, 1954)

PUNISHMENT: II. AN INTERPRETATION OF EMPIRICAL FINDINGS

JAMES A. DINSMOOR

Indiana University

In my first article on punishment (8), I pointed out that the procedures used in avoidance training and in punishment were alike in their basic pattern, differing only in relatively minor details. This being the case, I suggested that the principles we use to describe the conditioning and extinction of avoiding behavior might readily be extended to cover the operation of punishment. In other words, I suggested that the suppressive action of punishment was due to the conditioning of avoiding reactions which conflicted with the original behavior being punished. In this article I will try to fit such a description to the empirical findings, to see how well the two match.

The avoidance hypothesis can be stated in a little more detail. If we punish the subject for making a given response or sequence of responses—that is, apply aversive stimulation, like shock—the cues or discriminative stimuli for this response will correspond to the warning signals that are typically used in more direct studies of avoidance training. By his own response to these stimuli, the subject himself produces the punishing stimulus and pairs or correlates it with these signals. As a result, they too become aversive. In the meantime, any variations in the subject's behavior that interfere or conflict with the chain of reactions leading to the punishment delay the occurrence of the final response and the receipt of the stimulation that follows it. These variations in behavior disrupt the discriminative stimulus pattern for the continuation of the punished chain, changing the current stimulation from an aversive to

a nonaversive compound; they are conditioned, differentiated, and maintained by the reinforcing effects of this change in stimulation. The two types of behavior will soon reach a balance in strength, at a point where an occasional completion of the chain exposes the animal to a repetition of the punishment and this serves to maintain the avoiding reactions. Ending the punishment tends to lead toward a restoration of the original chain.

HISTORICAL OBJECTIONS

There are a few studies which we are forced to review merely because they have been widely cited in the past as *objections* to a systematic analysis of punishment. These are the "paradoxes" of the punishment literature, the experiments that seem to come out the wrong way, contrary to common sense. But when examined more closely, the problems turn out to be mainly problems of definition. Thorndike and his colleagues, for example, have conducted a number of studies (summarized in 36, 37) showing that small shocks or fines or the word "wrong" do not serve as very effective suppressors of behavior. This certainly should serve as a warning against too hastily extending general behavioral principles to complex social situations or against making careless assumptions about the functional role of stimuli we have not checked. But to conclude that *punishment* is not very effective raises the critical issue of just what we mean by "punishment" in the first place. How do we know these are punishments? Is there any way of

identifying an event as punishing other than by finding its effects on behavior? It is the action of stimuli that do suppress behavior that we must explain; if other stimuli have no effect on behavior, there is so much the less to explain.¹

Similarly, Tolman, Hall, and Bretnall (39) presented some human subjects with a stylus-punchboard maze and required them to choose one of each successive pair of holes. The subjects who were told that the *shocked* member of each pair was correct learned their route as well as or better than subjects who were told that the *nonshocked* member was correct. Normally, it is true, in the absence of special instructions, human subjects avoid shock; this is what led the authors of the study to speak of shock as a "punishment." But in this case the subjects were told *not* to avoid the shock, that the shocked response was the *correct* response; with these instructions, they no longer avoided it, but chose it. The shock no longer functioned as a punishment. What this study demonstrates is that the effects of aversive stimuli on human subjects—if they are not stronger than they "can stand"—can be overcome and obscured by appropriate verbal stimulation. The functioning of instructional stimuli may be an interesting and important problem in its own right, but it has little or no bearing on our attempt to analyze the action of punishment.

Finally, Muenzinger and his students (e.g., 26, 27) have shown that the effects of shock on behavior may be altered by still other procedures. For example, if rats are subjected to a mild shock each time they make the correct turn in a T-shaped discrimination apparatus, they learn the discrimination *more* quickly than if no shock is provided. This may seem paradoxical, but

one reason, at least, for this is that the shock is selectively paired or correlated with the food given as a reward for making the correct response. (Indeed, it is the reward that makes the response "correct.") As a result of this association, the shock becomes an immediate secondary reinforcer for the correct turn (9).

Another procedure Muenzinger tried was that of administering shock in *both* arms of the T. This also facilitates discrimination learning (13, 27). When retracing is allowed after incorrect entries, as in Muenzinger's work, the rat is reinforced by escape from shock for running toward the positive stimulus, no matter which arm he enters. When no correction is allowed, as in work by Freeburne and Taylor (13), the only remaining benefit may be that suggested by Muenzinger himself: that the animal proceeds more slowly through the choice point, with more head movements, and is exposed for a longer time to the discriminative stimuli. Note that these explanations supplement, but do not contradict, the avoidance hypothesis in a complex situation, and that they are required only to handle the "paradoxical" case.

PUNISHMENT AND ALTERNATIVE RESPONSES

We find one of the clearest and simplest illustrations of the complementary relationship between punishment and avoidance training in a situation provided by Warden and Aylesworth (40, punishment group). Their rats were let into the "reaction compartment" of a discrimination apparatus, where they faced another pair of compartments, one lighted and the other dark. On a given trial, the rat could do one of three things: (a) He could enter the dark compartment—this response was pun-

¹Thorndike's animal studies (35) illustrate extinction rather than punishment.

ished with shock, and he was then removed from the apparatus. (b) He could enter the lighted compartment—this allowed him to escape from the apparatus without being shocked. Or (c) he could refuse to make either response—after five minutes in the reaction compartment he would be removed by the experimenter.

Now consider the animals' behavior. If we assumed that punishment merely reduces the strength or probability of the punished response, leaving other behavior unaffected, we would predict that the animals would merely stop entering the dark compartment—that is, learn to remain in the reaction compartment. Again, if we assumed that punishment establishes a conditioned emotional reaction ("anxiety") which would reduce all of the animals' normal behavior (11), we would predict that they would stop entering either the lighted or the dark compartment; that is, they would become inactive and would remain in the reaction compartment. Indeed, this is what they *first* learned to do. By their second day of training, the rats used by Warden and Aylesworth responded this way on 80 per cent of their trials. But on later days, as they learned to discriminate between the dark compartment and the lighted one, they adopted more effective avoiding behavior, which led them more quickly out of the experimental situation: they learned to run into the lighted compartment.

The punished response does not simply dissipate, leaving the animal inactive; it is displaced by a response which increases in absolute as well as relative strength, at the expense of inactivity, as the training continues. Similar situations have been studied by Brown (3) and Robinson (28), with comparable results. The complementary relationship between the punished response and recorded alternatives has also been il-

lustrated in the bar-pressing situation by Edwards (10) and Sidman (32).

PUNISHMENT IN THE RUNWAY

When a rat is reinforced for running down an experimental alley or runway, the behavioral sequence is laid out for inspection along a single dimension in space. The successive steps that the animal takes in running down the alley bring him closer and closer to the further end; similarly, any avoiding reactions necessarily stop or reverse this movement. Miller (20) has shown that the net effect of first rewarding and then punishing the rat for approaching one end of an alley can be predicted from independent studies of the strength of the approach response and the strength of the avoiding response at different points along its length. This is very helpful in visualizing what happens when punishment is applied.

In some respects, however, this situation is not entirely typical. The successive members of the running chain are fairly homogeneous. One step is much like another. This makes it possible to treat them all, as Miller does, simply as successive instances of a single class of behavior—to speak of "the running response," "approach," or "locomotion." In like manner, both the proprioceptive stimulation and the external stimulation arising at different points along the runway remain very similar. Those changes in stimulation which do appear show no marked discontinuities or reversals in their progression from one end to the other. This enables Miller, for example, to treat the corresponding changes in the rate of running or the strength of pull at various points along the alley as the expression of a spatial gradient of stimulus generalization. Since this relationship tends to mask the usual chaining relationship, we should be fairly cautious in applying runway findings to other situations.

But the tendency for responses to generalize from one part of the runway to another also highlights certain relationships which we might not detect quite as readily in other experimental situations. In a pair of related studies carried out by Gwinn (15) and Whiteis (reported in 23, pp. 260-261), some rats were first trained to run down an alley to escape from a shock applied through the floor. When this training was completed, the procedure was changed. The charge was removed from the floor in the first portion of the alley but was retained in a later section: this meant that on a given run the animal first produced and then terminated the punishment by continuing forward. What apparently happened as a result was that the effects of the continued escape training in the charged part of the alley generalized to the earlier, uncharged portion. This served to maintain the animal's running and to suppress the avoiding reactions which would otherwise have appeared. Similar results have been obtained in direct studies of avoidance training. Mowrer and Lamoreaux (23, 24), for example, found that allowing the animal to turn off the shock, once it had occurred, by means of a response which was different from the avoiding response, led to slower avoidance learning; allowing the animal to turn off the shock by means of the same response facilitated avoidance learning.

Also: since the running chain is so homogeneous from one point to the next and provides so little differential stimulation, special external stimuli may have to be furnished to provide adequate secondary stimuli or warning signals for the avoiding responses (23, p. 261). Analogous data showing that the efficiency of formal avoidance training likewise depends on the distinctiveness of the warning signal have been reported

by Kessen (18), Miller and Greene (22), and Mowrer and Lamoreaux (25).

THE FREE RESPONDING SITUATION

In highly restricted situations like the discrimination apparatus and the runway, where we measure the relative strength of two or more alternative responses, the relation between avoidance training and punishment might seem to be "built in"; as one response decreases in strength, another must increase, and as one increases, another must decrease. Can we extend the application of the avoidance hypothesis to the more general case where the *absolute* frequency or rate of responding is recorded? Here we must study the changes in rate following upon a variety of experimental operations and see whether or not they fit our description. There are three stages in the application of punishment at which we can test deductions from an avoidance hypothesis: (a) We can see how quickly the response is suppressed when the punishment is first administered. (b) We can see what happens to the rate of responding as the punishment continues. (c) We can see what happens to the response when punishment is no longer applied. (In each case, our deduction seems to be confirmed.)

SPEED OF LEARNING

In a free responding situation, the initial change in the rate of responding should be relatively rapid when punishment is first applied, more rapid than the changes observed in a direct study of avoidance training. This conclusion follows from a consideration of some of the differences between the two procedures reviewed in my previous article (8).

In avoidance training, the animal is required to meet certain narrow and

arbitrary specifications of his behavior to avoid the punishment. The frequency of this narrow class of behavior prior to the training is likely to be relatively low. And no opportunity is provided for the selective strengthening (i.e., "shaping up" or differentiation) of this particular form of behavior from among a set of partly successful variations, since nothing less than an actual depression of the bar, for example, will serve to postpone the shock. This would tend to slow down the rate of conditioning. Also, the alternative behavior to the avoiding response is quite varied, which means that a large number of applications of shock may be necessary before the stimuli accompanying these varied responses have all been made aversive to the subject. Only then does the execution of the correct response consistently lead to a reinforcing change in stimulation.

But when a punishing procedure is used, there is only one chain of responses that leads to aversive stimulation, and the accompanying stimuli are followed fairly regularly by the punishment. Alternative behavior begins to receive reinforcement almost immediately. Note also that there are many possible forms of avoiding behavior. This means that their collective frequency should be relatively high at the beginning of the experiment and that more successful forms may be expected to emerge from the less successful as a function of variations in the length of the delay they produce in the appearance of the aversive stimulus.

As yet, we have no rigorously controlled test of this prediction, but a more casual comparison suggests that the deduction is correct. Avoidance studies in a free responding situation (e.g., 23, 30) indicate a relatively slow process; in studies of punishment the suppression is almost immediate (7, 10, 11, 33).

BEHAVIOR DURING PUNISHMENT

There is a limiting factor in the frequency with which the animal will make an avoiding response. As the animal continues to perform the required response over and over again, the efficacy of the stimulus change produced by this response gradually declines, since the secondary stimulation (e.g., warning signal) is no longer followed by the shock. This trend is self-correcting, however, for when the animal eventually fails to respond, he permits a repetition of the original pairing between the two stimuli, and this leads to intermittent reconditioning. The frequency of response should fluctuate within circumscribed limits, and a relatively stable rate of responding should and does emerge (31).

When the avoiding response is pitted against another response that has been maintained at a fairly stable rate, as in the case of intermittently reinforced bar pressing, the two forms of behavior should soon reach an equilibrium (7; 10; 11, Experiment H). As Estes has put it, "Punishment of a periodically reinforced response results simply in an adjustment of the rate of responding to a lower, but still constant, value which is determined jointly by the conditions of reinforcement and the severity of the punishment" (11, p. 23). When no reinforcement is provided for the punished response, its rate shows a slow but steady decline (7, 10).

RECOVERY FROM PUNISHMENT: "COMPENSATION"

On the basis of an avoidance interpretation, we would expect the animal to continue to respond at a reduced rate for some time after we have stopped punishing him. The avoiding responses should persist until the animal has received sufficient exposure to the experimental situation and to the stimulation produced by his own behavior for these

stimulus patterns to lose the aversive properties they acquired while they were being paired with the punishment. Less and less effective reinforcement would then be provided for the avoiding reactions. These should extinguish, and the animal should gradually return to some extent toward his original mode of behavior.

The length of time required for this recovery and the level of responding which eventually is attained should depend on such factors as the number and schedule of previous reinforcements, the duration of the punishment procedure, and the intensity of the punishing stimulus. At the one extreme, with a strong response and with brief and mild punishment, the animal should return approximately to his original rate of responding, since the chief effect of the punishment has been simply to prevent the occurrence of the response for a time by blocking or interference (11, Experiment H). If no reinforcement has been provided during and after the punishment, he should reach a peak which is slightly lower than his original rate but which is higher than the current rate of responding among control animals (who have been extinguished but not punished during the same period). And he might be expected to have made about as many responses as these control animals by the time extinction has been completed (11, Experiment A; 33, pp. 154-155).

It is difficult to see how we could account for this "catching up" with the control totals on the basis of any formulation which does not involve the concept of the animal's behavior being blocked by competing responses. It is characteristic of situations where the bar is removed for a time and the animal prevented from responding (4). It is not characteristic of procedures where the animal's level of activity is reduced

by the provision of food or drugs (e.g., 5). In particular, it is not deducible from the postulate of a "depressed state of the organism" such as Estes' anxiety state (11), unless a subsequent emotional state of "joy" or "relief" is postulated, *ad hoc*, or the anxiety state itself is to be characterized as consisting of interfering behavior of some kind.

EXTINCTION DURING PUNISHMENT

When the animal has been punished for a longer time, or when the avoiding reactions have persisted for some time owing to the severity of the punishment, a reduction in the animal's resistance to extinction should become detectable. That is, the punished animals should fail to respond as many times during extinction as their nonpunished controls. In deducing this effect, we must consider the entire chain of stimuli and responses leading up to the punishment.

Skinner (33) has analyzed the bar-pressing chain, for example, into a series of behavioral components, such as approaching the bar, raising the paws, making contact with the bar, pressing, releasing, approaching the food tray, seizing the pellet, and so on. Each of these responses produces one or more discriminative stimuli for the next response in the sequence. By selecting a stimulus that he could present or withhold at will (e.g., sound of food magazine operating, delivery of pellet) and omitting it from the procedure, Skinner found that he could extinguish the responses which *preceded* this stimulus in the chain. Responses which *followed* this stimulus in the chain, however, did not seem to have been affected: when the missing stimulus was restored to the experimental procedure, these responses immediately reappeared. Restoration of the sound of the food magazine, for example, immediately led to approaches to the food tray, and restoration of the

pellet immediately led to eating (33, pp. 102 ff.).

But in Skinner's procedure the chain was broken between a given *response* and the ensuing *stimulus*; in punishment, the chain may be broken at a number of places by interfering behavior, but in each case the break is between a given *stimulus* and the following *response*. Thus, the animal may carry out the first part of the sequence and may produce the discriminative stimuli for the punished response. These stimuli exercise their customary reinforcing influence on the responses that have already appeared, but avoiding reactions prevent the occurrence of the punished response itself. Nothing shows up on the record. But according to a previous analysis I have presented (6), the use of a stimulus as a reinforcer subtracts from its future effectiveness as a discriminative stimulus and thereby from the rate of occurrence and resistance to extinction to be expected for the next response in the chain. "The extinction of a sequence $R \rightarrow S$ " may reduce the strength of the following sequence $S^0 \cdot R$ " (6, p. 471).

If, then, the animal is led to continue to make avoiding responses for a long enough period, we can predict: (a) that the peak rate of responding following the punishment may be no greater than that for unpunished control animals at the same point in the experiment; and (b) that the number of recorded responses during extinction will be significantly reduced by a period of punishment. This has been repeatedly confirmed (7; 10; 11, Experiments B through G).

ANXIETY AND PUNISHMENT

We have reviewed the major changes that take place in the rate and resistance to extinction of a punished response and we have seen that they can

be deduced from the assumption that the response is suppressed by competing avoidance behavior. The remaining details also seem to fit in quite readily with this interpretation. But in his monograph, Estes has suggested another way of dealing with these results: he has postulated the conditioning of an emotional state, "anxiety," which is said to depress the normal behavior of the subject (11).

Estes' account becomes a bit arbitrary in spots, for some of the characteristics he has assigned to the anxiety state seem to have been derived from the very data that they are designed to explain. This device may have been useful as a first approximation, to summarize the original data. But with one exception, it seems to be difficult to extend the anxiety concept beyond the original data; it does not help us, for example, to explain the animal's behavior in a runway or a discrimination apparatus. In general, its power to handle additional data seems to be quite limited. By contrast, the avoidance hypothesis seems to be much more efficient, since it provides a unified treatment, on the basis of limited assumptions, for a much greater variety of data.

There is one further area of investigation, however, which fits rather neatly into Estes' anxiety formulation and which at first might seem to pose some difficulties for the avoidance hypothesis. The aversive stimulus need not be paired with the animal's response to produce its suppression: pulses of shock may be interspersed between responses (11, Experiment I), or a neutral stimulus may be paired with shock, on an arbitrary schedule (e.g., 1, 2, 12, 14, 17, 19). Again, we could appeal to an emotional state to account for the observed changes in the organization of the bar-pressing sequence (14) or for the reduction in its frequency of occurrence (1, 2, 11, 12, 17, 19). But the necessity for this ap-

peal is again questionable, since there are observable reactions by the subject which may be sufficient to account for the changes in the recorded response.

To begin with, there are a number of skeletal muscle reactions to the first application of a shock, such as jumping, running, crouching, or biting the grid. These may be escape reactions which have generalized from other forms of painful stimulation, or they may be innate reactions to shock as an unconditioned stimulus. In either case, we can expect some transfer of these reactions to the secondary or warning stimuli that accompany the shock. And they would conflict with other skeletal muscle behavior.

Another likely source of skeletal muscle interference is the type of avoidance learning described by Hefferline (16, Experiment 3), Tolcott (38), and Winnick (41). In these studies, a bright light, rather than a shock, was used as the aversive stimulus. The light was presented continuously, except when the animal adopted a specified position or response, such as stepping on a pedal or pushing in a wall panel with the nose. Under these conditions the rat learns a prolonged or sustained "holding" response, which interferes with most other forms of activity. But there is no indication of any reduction in activity, as would be implied by the anxiety hypothesis, when the animal lets go and actually exposes himself to the light. On the contrary, random activity seems to increase (16), and experimentally reinforced activity definitely increases above its normal rate (38, 41) in intervals between the holding behavior.

Similar avoiding behavior probably becomes conditioned in the presence of any continued aversive stimulus, whether it is the primary stimulus itself, such as light or shock, or a stimulus which has been paired with the primary stimulus. A number of writers have commented

on the many "ingenious" ways which the animal finds to avoid or reduce shock or light stimulation, and similar devices can probably be found for auditory stimulation. In one case, at least, experimental arrangements designed to reduce or eliminate such devices also eliminate the usual intermittent depressions in the rate at which the animal makes the recorded response (34). Whether we can ever prove or disprove the necessity for supplementing this explanation with an appeal to some form of emotional state remains to be seen; but for the present, the burden of proof seems to rest on those who do maintain the necessity of such an appeal.

Other writers have proposed another form of fear or anxiety, which is said to operate in avoidance and in punishment as an "acquired drive" (e.g., 21, 23). None of these writers, however, have made it at all clear what they mean by a drive in the first place, innate or acquired. If a drive is that which is reduced when reinforcement is observed, and no more, then the term is merely a second name for a single event (see 29). If a drive is a prior condition that permits reinforcement—e.g., the presence of shock permitting the reduction or termination of shock—then the concept remains quite trivial. It is not itself a condition governing behavior but a physical precondition for an operation which is already well recognized and which has its own label. It points out no additional variable for the prediction or control of behavior. If the drive of fear or anxiety, in particular, is assumed to have functional properties quite the opposite of those assumed by Estes—that is, to *multiply* habit strength—then the positive evidence (17a) is as yet rather scanty. Even as a device for the popularization of psychological principles, the concept has its drawbacks. What we call fear in everyday situations seems to be a mode of reaction (e.g.,

withdrawing, hiding, appealing) rather than a physiological or hypothetical state. And while avoiding behavior may sometimes be classified as fear, other forms are frequently classified under the heading of anger or a number of other emotions.

SUMMARY

In this article I have applied the avoidance interpretation of punishment to a variety of empirical findings. I have tried to show that some of the studies frequently cited as objections to a systematic account are actually irrelevant to explaining the effects of punishment. I have reviewed the interaction of punished responses and avoiding behavior in the discrimination apparatus and the runway and have deduced the observed patterns of change in the frequency of free operants from this hypothesis. On the basis of this review, I would conclude that this hypothesis provides an adequate account for the main facts of punishment. No appeal to a state or drive of anxiety as yet seems to be required.

REFERENCES

1. AMSEL, A. The effect upon level of consummatory response of the addition of anxiety to a motivational complex. *J. exp. Psychol.*, 1950, **40**, 709-715.
2. AMSEL, A., & COLE, K. F. Generalization of fear-motivated interference with water intake. *J. exp. Psychol.*, 1953, **46**, 243-247.
3. BROWN, J. S. Factors determining conflict reactions in difficult discriminations. *J. exp. Psychol.*, 1942, **31**, 272-292.
4. BULLOCK, D. H. Operant extinction as a function of the extinction schedule. *J. exp. Psychol.*, 1951, **42**, 437-441.
5. CROCKETT, C. P. The relation of extinction responding to drive level in the white rat. Unpublished doctor's dissertation, Columbia Univer., 1951.
6. DINSMOOR, J. A. A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp. Psychol.*, 1950, **40**, 458-472.
7. DINSMOOR, J. A. A discrimination based on punishment. *Quart. J. exp. Psychol.*, 1952, **4**, 27-45.
8. DINSMOOR, J. A. Punishment: I. The avoidance hypothesis. *Psychol. Rev.*, 1954, **61**, 34-46.
9. DINSMOOR, J. A., JOHNS, G. R., KENT, N. D., SIMON, W. B., & WINDMAN, GEORGIA O. Electric shock as a secondary reinforcer for the correct response in Muenzinger's studies of discrimination learning. Paper read at Midwest. Psychol. Ass., Columbus, Ohio, April, 1954.
10. EDWARDS, RUTH P. The effects of punishment on unpunished responses. Unpublished doctor's dissertation, Harvard Univer., 1951.
11. ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**, No. 3 (Whole No. 263).
12. ESTES, W. K., & SKINNER, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, **29**, 390-400.
13. FREEBURNE, C. M., & TAYLOR, J. E. Discrimination learning with shock for right and wrong responses in the same subjects. *J. comp. physiol. Psychol.*, 1952, **45**, 264-268.
14. FRICK, F. C. The effect of anxiety—a problem in measurement. *J. comp. physiol. Psychol.*, 1953, **46**, 120-123.
15. GWINN, G. T. The effects of punishment on acts motivated by fear. *J. exp. Psychol.*, 1949, **39**, 260-269.
16. HEFFERLINE, R. F. An experimental study of avoidance. *Genet. Psychol. Monogr.*, 1950, **42**, 231-334.
17. HUNT, H. F., & BRADY, J. V. Some effects of electro-convulsive shock on a conditioned emotional response ("anxiety"). *J. comp. physiol. Psychol.*, 1951, **44**, 88-98.
- 17a. KALISH, H. I. Strength of fear as a function of the number of acquisition and extinction trials. *J. exp. Psychol.*, 1954, **47**, 1-9.
18. KESSEN, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45**, 82-86.
19. LIBBY, A. Two variables in the acquisition of depressant properties by a stimulus. *J. exp. Psychol.*, 1951, **42**, 100-107.
20. MILLER, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. 1. New York: Ronald, 1944. Pp. 431-465.
21. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Hand-*

- book of experimental psychology. New York: Wiley, 1951. Pp. 435-472.
22. MILLER, W. C., & GREENE, J. E. Generalization of an avoidance response to various intensities of sound. *J. comp. physiol. Psychol.*, 1954, **47**, 136-139.
 23. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
 24. MOWRER, O. H., & LAMOREAUX, R. R. Fear as an intervening variable in avoidance conditioning. *J. comp. Psychol.*, 1946, **39**, 29-50.
 25. MOWRER, O. H., & LAMOREAUX, R. R. Conditioning and conditionality (discrimination). *Psychol. Rev.*, 1951, **58**, 196-212.
 26. MUENZINGER, K. F. Motivation in learning. I. Electric shock for correct response in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267-277.
 27. MUENZINGER, K. F., BERNSTONE, A. H., & RICHARDS, L. Motivation in learning. VIII. Equivalent amounts of electric shock for right and wrong responses in a visual discrimination habit. *J. comp. Psychol.*, 1938, **26**, 177-186.
 28. ROBINSON, J. S. Stimulus substitution and response learning in the earthworm. *J. comp. physiol. Psychol.*, 1953, **46**, 262-266.
 29. SCHOENFELD, W. N. An experimental approach to anxiety, escape, and avoidance behavior. In P. J. Hoch & J. Zubin (Eds.), *Anxiety*. New York: Grune & Stratton, 1950. Pp. 70-99.
 30. SIDMAN, M. Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science*, 1953, **116**, 157-158.
 31. SIDMAN, M. Two temporal parameters of the maintenance of avoidance behavior by the white rat. *J. comp. physiol. Psychol.*, 1953, **46**, 253-261.
 32. SIDMAN, M. Delayed-punishment effects mediated by competing behavior. *J. comp. physiol. Psychol.*, 1954, **47**, 145-147.
 33. SKINNER, B. F. *The behavior of organisms*. New York: D. Appleton-Century, 1938.
 34. SKINNER, B. F., & CAMPBELL, S. L. An automatic shocking-grid apparatus for continuous use. *J. comp. physiol. Psychol.*, 1947, **40**, 305-307.
 35. THORNDIKE, E. L. Rewards and punishment in animal learning. *Comp. Psychol. Monogr.*, 1932, **8**, No. 4 (Whole No. 39).
 36. THORNDIKE, E. L. *The fundamentals of learning*. New York: Bureau of Publications, Teachers Coll., 1932.
 37. THORNDIKE, E. L. *The psychology of wants, interests, and attitudes*. New York: Appleton-Century, 1935.
 38. TOLCOTT, M. A. Conflict: a study of some interactions between appetite and aversion in the white rat. *Genet. Psychol. Monogr.*, 1948, **38**, 83-142.
 39. TOLMAN, E. C., HALL, C. S., & BRETNALL, E. P. A disproof of the Law of Effect and a substitution of the Laws of Emphasis, Motivation, and Disruption. *J. exp. Psychol.*, 1932, **15**, 601-614.
 40. WARDEN, C. J., & AYLESWORTH, MERCY. The relative value of reward and punishment in the formation of a visual discrimination habit in the white rat. *J. comp. Psychol.*, 1927, **7**, 117-127.
 41. WINNICK, WILMA A. A study of incipient movements in avoidance. Unpublished doctor's dissertation, Columbia Univer., 1950.

(Received November 18, 1953)

THE DYNAMICS OF IDENTIFICATION¹

NEVITT SANFORD

University of California, Berkeley, and Mary Conover Mellon Foundation, Vassar College

R. P. Knight, writing in 1940 (11), stated that "identification" was probably used in more different ways than any other psychoanalytic term. There is nothing to suggest that the term has been pinned down since then. Knight was concerned with uses by psychoanalytic writers; other psychologists and social scientists have, it seems, found additional ways in which the term could render service.

The following passage from Tolman (18) shows something of the extent of our problem; it also calls attention to an interesting bit of history. "Identification was apparently first noted and named by Freud. But his conception became unnecessarily complicated and it was too closely bound up with his whole psychoanalytical system. I shall not mean here by identification, therefore, Freud's own concept, but merely a certain general neo-Freudian notion which seems now to be widely accepted by most psychologists and sociologists" (18, p. 141). Tolman then goes on to mention three different, though related, processes covered by his general neo-Freudian notion. (a) "The process wherein an individual tries to copy—to take as his pattern or model—some other older (or in some other way looked-up-to or envied) individual"; (b) "The adherence of the individual to any group of which he feels himself a part"; and (c) "The acceptance by an individual of a cause" (18, pp. 141-142).

¹ This paper was a contribution to the symposium "The Identification Concept and the Theory of Personality and Psychopathology" (Joseph Adelson, chairman) held at the meeting of the American Psychological Association, Cleveland, 1953.

It may be added that the term identification is also commonly used to refer to the phenomena of empathy and of vicarious living, of sympathy and altruism, and that it creeps into our vocabulary when we try to describe closeness, or loyalty, or even conformity or submissiveness, as between two people. (And, furthermore, we know that the objects of identification, as the term is variously used, are not confined to other people, singly or in groups, but may include animals, machines, inanimate objects, parts or features of people; and that identification may be expressed not alone in overt behavior but in conscious experience, in attitude, in fantasy.)

More than this, identification is used not only in the description of a broad area of everyday behavior; it frequently refers to a mechanism or process by which the personality is changed. Mowrer, in his recent work (13), distinguishes between "developmental identification," in which the child *learns* to perform ego functions like his parents, and "defensive identification," in which the child accepts the standards of his parents as a means for pleasing them and as a means for controlling his own impulses. (By way of entering into the spirit of the thing, I am going to suggest, later, that both of the processes referred to by Mowrer are developmental and that neither is properly called identification.)

In much classical psychoanalytic literature "introjection" is the word for that process, or those processes, by which the individual takes over, and makes his own, psychological attributes of other people, but we find that identification is used interchangeably with or in place of introjection, as well as to

stand for a different process. More recently Freudian psychoanalysts, in turning their attention to everyday social behavior, have not hesitated to use the term identification in much the same way that Tolman does, that is, to stand for certain activities of the conscious ego and for certain common patterns of everyday social behavior. Thus, for example, Alexander speaks of "the introspective knowledge of one's own emotions which one uses through identification in the understanding of others" (2, p. 41); Balint (3) coins the term "genital identification" to stand for that kind of mutual understanding that we find in a mature love relationship. And Reider (16) leans heavily upon the concept of identification in explaining good morale in organizations of the armed services.

I have a theory about what has happened. The manifest phenomena of identification, that is to say, the kinds of social behavior mentioned above, were observed before Freud and attempts at their description have gone forward independently of psychoanalysis—as well they might. But psychoanalysis has been in the air; and psychologists like Tolman and Mowrer, with systems of their own, have from time to time borrowed psychoanalytic terms to stand for conscious ego processes, or aspects of animal behavior, that resembled what Freud had in mind only in some loosely analogous way. Psychoanalysts, for their part, as they have become increasingly interested in everyday matters and increasingly concerned about Freud's neglect of ego psychology, have tended to apply to these surface phenomena concepts originally designed for what was primitive, infantile, unconscious. Thus, the watering down of Freudian concepts, it must be admitted, is to be ascribed mainly to Freudian psychoanalysts. Apparently, out of a need to show that Freud's concepts were adequate for the whole range of human

behavior, there has been a tendency to stretch and dilute them, sometimes almost beyond recognition. We are not misled by Tolman and Mowrer because, though we expect much of them, we do not expect them to be more psychoanalytic than Freud; from psychoanalysts we expect to hear about psychoanalysis, and are unprepared to see an everyday social phenomenon treated as if it were a depth-psychological one. One might suggest that as academic psychologists and psychoanalysts have glowered at each other through the years there has occurred a certain amount of "identification with the enemy," so that increased similarity of the two groups may be noted; unfortunately, however, new patterns acquired in this way tend not to be integrated with the ego; in so far as they persist, it is as "foreign bodies" within the personality.

A term that can be employed in so many different ways and that, as Tolman says, has been accepted by most psychologists and sociologists, could hardly mean anything very precise. It might be proposed, quite seriously, that we give up the term "identification" altogether. When we are describing social behavior and have the impulse to say "identification" we must in any case specify "what kind"; if we go a step further and say just what we mean, it will almost always turn out that other words are available, and that they are, in fact, more accurate. And when it comes to explaining behavior we might agree with Knight that the phenomena ordinarily called identification are "always based on a subtle interaction of both introjective and projective mechanisms" (11, p. 335). In other words, why not agree that identification is not an explanatory concept, and that as a descriptive one it is too vague to be useful? A moratorium on identification would not, as one might suspect, leave the clinical psychologist quite inarticulate.

late. But it might well sharpen his observations of behavior and challenge some of his theoretical assumptions.

I have tried the experiment recently—the experiment of doing without identification—in organizing material on therapeutic cases. The result was that in formulating the dynamics of the case and in writing out the developmental history, one could do very nicely without this term. For describing the common social relationships to which the term has been applied, such words as love, friendship, closeness, loyalty, alliance, solidarity, empathy, fellow feeling, kinship, understanding, sympathy, participation, vicarious living, submission, and acceptance seemed fairly adequate, and in the consideration of deeper determining factors one could lean, rather heavily to be sure, on introjection, and on learning and various kinds of constructive ego functioning. But—and this is for me a very significant fact—when it came to dealing with certain aspects of the changing therapeutic relationship the concept of identification seemed indispensable. (Naturally, what is true of psychotherapy would be true of other interpersonal relationships.)

I do not mean my identification with the patients; I mean their identification with me. Let it be granted that living vicariously in one's patients, getting emotional excitement from their recountings of all the daring and fantastic and all-too-human things that they do, and that we do not have time for, can be quite satisfying (the therapist does not need to read novels or watch television); and let it be insisted that the only helpful understanding of a patient rests upon the therapist's ability to put himself emotionally in that patient's place. Both of these processes, I believe we may hope, depend upon broadness of consciousness, grasp of reality, sureness of self-control; they are of a piece with the ability to love, and to experience a

sense of kinship with one's fellow man. When a patient may properly be said to identify with the therapist, on the other hand, we deal with a process that is unconscious and unrealistic, with a patient who is unsure of himself and, at the moment at least, unconcerned about other people; in desperation he adopts a piece of poor economy as a means of escape from a critical situation. I believe that it is stretching the term identification too much to apply it both to what happens in the therapist—when he is functioning well—and to what happens in the patient; such usage may easily tend to obscure both processes.

We are, of course, accustomed to the idea that the same behavior, or very similar patterns of behavior, may have quite different sources. No one can deny the merit of Knight's suggestion that with respect to the phenomena loosely called identification we make a sharp distinction between source and surface and proceed with the business of explanation. But I am proposing that phenomena have been grouped under the heading of identification that are not similar even in their most manifest aspects, or at the least that this similarity is insignificant as compared with their differences. The patient's identification with the therapist and the therapist's understanding of and vicarious satisfaction through the patient constitute a case in point, and others will be offered below.

I am also going to propose that these two phenomena have quite different determinants. Disagreeing with Knight, I would argue that the therapist's constructive and enjoyable reactions to the patient have little if anything to do with introjection or projection, beyond sometimes being in some general way analogous to them. I am also going to propose that identification, in the sense in which I shall define it, is not a category of behavior but a "mechanism." As a

"mechanism," it can be distinguished from introjection, and from other processes by which features of the environment are taken into the personality.

In identification proper, that is, in identification as here understood, the individual may be observed to respond to the behavior of other people or objects by initiating in fantasy or in reality the same behavior himself. This is identification of the self with the object; it is different from empathy, fellow feeling, vicarious living, and the like—those phenomena which Knight has properly called identification of the object with the self.

Identification proper is unconscious, or at least more or less unconscious. This differentiates it from conscious imitation, and other processes by which we more or less accept other people's ways of doing things because we find that they serve us well. When we observe true identification in our patients we interpret it; and they characteristically become aware of it with a measure of embarrassment.

Most important, perhaps, identification tends to be *identical*, that is to say, the subject strives to behave in a way that is *exactly* like that of the object. We may note the identity, or the attempt at identicalness, in *detail*, if not in many details at least in *concrete* aspects of the object's behavior. It is this that permits us to place identification proper in the same class with those forms of behavior that tend to be rigid, all out, total—with those reactions which are switched in when, in a critical situation, the limits of the individual's soundly economic modes of adjustment are surpassed. Thus we may usually note a measure of desperation in the subject who identifies. I say this feature of identification proper is most important because it really provides us with the key to its understanding; it is the touchstone of its dynamic nature;

identification is one of those reactions—adaptive in the short run, maladaptive in the long run—to which the individual, any individual, may be driven by circumstances.

Identification as here understood is unrealistic. Although the subject strives to be exactly like the object in some respect or altogether, it is rare that he can be in any real sense. We may note striking similarities, but we may also note an aspect of caricature. Behavior in identification is likely to be mechanical or otherwise incongruous; since the subject is not being himself, the behavior that we observe is likely to appear as foreign to him. And, finally, the behavior is rarely, if ever, an appropriate means to any end that is in the subject's long-term interest.

In identification proper, the object of identification is *there*, a part of the subject's external situation. Usually the subject, in identifying, not only perceives the object at the time but is, or can be, perceived by the object. Frequently, it is important to him that his identical behavior be perceived by the object, or by other people who have some importance in the general situation.

In turning from manifestations of identification in behavior to a consideration of underlying dynamics, let us examine the case of the patient who identifies with the therapist, that is to say, who exhibits one kind of such identification, that which we understand as resistance. Probably most therapists can confirm the original observation of Abraham (1). In describing "A form of neurotic resistance" found predominantly in narcissistic patients, Abraham writes, "In place of making a transference, these patients tend to identify themselves with the physician. Instead of coming into close relation to him they put themselves in his place. They adopt his interests and like to occupy themselves with psychoanalysis as a science,

instead of allowing it to act upon them as a method of treatment. They tend to exchange parts, just as a child does when it plays at being father" (1, p. 306). What seems most apparent in such cases is that the subject tries, rather desperately, to protect his self-esteem. The phenomenon seems quite comparable to those most striking instances of identification in childhood: the little boy who must act just like father in some particular, the little girl who exhibits some of her mother's patterns of behavior in an incongruous, grown-up manner. This is by no means always play; the child is frequently deadly serious, and we receive a strong impression of his insecurity. It is as if he cannot be the child that he is, as if he cannot tolerate the sense of weakness or smallness or danger which he feels goes with that role, but must hurry and be, or act as if he were, something different.

This is not, in my opinion, the same process as that by which a child, in the usual case, slowly adopts some of the standards and ways of his admired and loved parents—adopts them and holds onto them because they serve his long-term needs or purposes. In this latter instance we are not likely to be struck by the identity of his behavior and that of the parent. What he adopts for himself, in contrast to what he merely borrows, he quickly puts his own stamp upon, and this from a very early age, so that when we compare him and his parents the most that can be noted is a general similarity. If I seem to be saying that the child who *really* identifies with his parents tends not to exhibit signs of identification, this is close to being just what I mean; but I must of course find other terms for the "real" identification, the integration of parental standards and modes of behavior in a stable ego system.

It would appear, then, that one dy-

namic source of identification proper is a threat to self-esteem, a threat that is severe enough so that realistic methods for coping with it do not suffice.

Threats to the subject's physical existence or to the physical integrity of the organism may also lead to identification. Apparently the aim of identification in such instances is to acquire a sense of power and hence to feel equal to the threat. (A two-year-old boy of my acquaintance was terrified by a new puppy that was brought into the house; but after a few hours this boy was crawling about making barking noises and threatening to bite people; his fear of the puppy had vanished.) Sometimes in situations of dire threat, the individual seeks to protect himself not so much by copying the aggressor as by going over to his side, by joining forces with him. This may be a conscious stratagem more or less deliberately, or cynically, chosen; but sometimes, the whole proceeding is largely unconscious, the subject deceiving himself in judging his own interests as well as the motivations and moral justification of his enemy or authority.

In cases of extreme domination, where the existence of the subject as an independent—choosing, decision-making—individual is threatened, it seems that he has the alternatives of either taking over the ways of his guard or parent or of having no personality at all. There may be no possible alternative but to submit, but in submitting one may still maintain some sense of self through participating in the personality of his oppressor. Members of ethnic or sexual minorities seem often to exhibit this kind of identification with the dominant group. A patient in extended therapy may show a similar reaction at a certain stage of the therapeutic relationship, e.g., when dependence on the therapist is intensified by real difficulties

and by the necessity of making numerous readjustments.

It would appear, then, that identification proper is a desperate attempt to deal with a crisis involving the self. We have seen that such a crisis is commonly precipitated by the aggressive or dominant actions of another person or object, in the subject's immediate environment, from whom no means of escape can be found. At the same time, it is important to note that the crisis may arise out of events that are chiefly internal, in which case the object of identification is likely to be chosen by the subject. An arousal of impulses belonging to a negative identity, for example, feminine impulses in a male, may lead to hasty attempts at identifications having an opposite direction and meaning. Generalized social insecurity, crises in identity like those so frequently endured by new students at boarding school or college, often lead to true identification with some of the common ways of a peer group.

It has been stated that when crises of selfhood have sources that are chiefly internal, the object of identification will be chosen by the subject. Now it must be added that some objects are much more likely to be chosen than others, that there are, indeed, "identification evokers"—heroic or sensational figures that seem able to elicit identification in almost anyone. The stronger the stimulus the less intense the inner crisis needs to be in order for identification to occur. In most people, probably, there is enough uncertainty of identity or dissatisfaction with the self so that they may be induced to take an ill-advised short cut to improvement. More than this, most clear-cut instances of identification with the dramatic figure are mass phenomena, or at least phenomena involving two or more identifiers. Once identification is made by one or two or three individuals in a group, presumably those in the

severest straits with respect to self-esteem, the situation changes; now it becomes difficult to tell whether copying the dramatic figure is induced primarily by *him* or primarily by social pressure. In either case, we have to deal with a regressive attempt either to enhance the self or to protect its existing status.

The present rather critical view of identification seems a necessary concomitant of an attempt to get the term back into its place, so to speak, partly by reviving some of its traditional meanings. Freud consistently spoke of identification as primitive or regressive—even when his use of the term was similar to the present one. It is only in comparatively recent years, since the onset of the tendency among psychologists and psychoanalysts to "accentuate the positive," that the concept has been given an important place in sound character development, successful psychotherapy, harmonious individual-group relations, and the like. But, since the phenomena of identification proper cannot be overlooked, it has become common practice among textbook writers to speak of "good" identifications and "bad" identifications, or of too much or too little. Symonds (17), for example, lists the "positive" and the "negative" values of identification. Some of the positive values listed, e.g., that it gives a person ambitions and ideals or supplies a basis of sympathy, should, according to the present view, be attributed to processes other than identification; the others, e.g., that it may help the individual gain security or strength, should be regarded as "good" only in the sense that they help stave off something worse or represent the best that is possible in the circumstances; as adjustments or adaptations made in an emergency, they serve to hold the fort, as it were, until preparations can be made for a genuine developmental advance.

Identification is undesirable for Sy-

monds when it is exaggerated or fantastic or without discrimination. This has been called by other writers *over-identification*. But when we are offered an example of the proper amount of identification, it usually turns out to be not a lesser degree of the same process, but a different process altogether—most commonly, perhaps, the learning of something true or useful from somebody else. The argument here is that we may reduce confusion by holding ourselves to the use of the term identification to stand only for the ultimately maladaptive crisis reaction.

The discussion so far has been concerned solely with behavior in a momentary situation. I have sought to view the phenomena of identification in field-theoretical terms, the hope being that this may be an aid to experimental attacks on the problem before us. It is commonly assumed, however, that identifications once made tend to persist, that the type of crisis reaction considered here results in lasting changes in the personality. It is a widely accepted theory that "defensive identification," or identification with the feared and frustrating parent of the same sex, is an important source of the superego. The remainder of this paper will argue, first, that the lasting effects of identification proper are frequently overestimated; second, that the process by which figures of the environment become models for response readinesses in the deeper layers of the personality, as in superego formation, is not identification but introjection; and finally, that there are other processes besides these depth-psychological ones by which environmental influences enter into the moulding of character.

If, to take up the first argument, identification proper is a function of a total situation and is to be understood in field-theoretical terms, then we should expect the identification behavior to

change as soon as the situation changes. It seems to me that in the usual case this is just what happens; identification is evoked by a crisis, the crisis passes or is mastered, and we are able to observe no persisting effects.

It should be borne in mind here that in the case of the resistant patient, identification with the therapist is a means of keeping him out, of not being influenced by him. I think Abraham is right in likening this behavior to a child's defiance of his parents. In identification with the aggressor the subject does not bring about a basic change in his relations with the threatening figure by his acts of identification, as would be the case should the aggressive pattern become integrated with the ego system; the aggressor remains outside, to be dealt with by repetitions of the same stratagem. The subject who "borrows another personality" in order to achieve a desperately needed identity does not solve his problem in this way; he is defeated either by reality or by psychosis; the adolescent who switches rapidly from one idealized object of identification to another does not finally hit upon one that suits him; he learns finally that he can only be himself. Much consistent behavior that appears to be an expression of an identification with some figure of the past can best be explained, it seems to me, on the basis that a device which proved successful on one occasion is likely to be repeated in an equivalent one. In the authoritarian personality the inclination to identify with figures of authority is, of course, marked, and this is readily traced back to early patterns of identification with parents; but the striking thing is that the authority has not been genuinely internalized, as we know from the fact that it does not operate effectively unless a representative of it is present or close by. We may speak here of an infantile or rudimentary superego, and

we are not surprised to find that identification with authority figures is the more marked the less the visible signs of a firmly established inner conscience. But actual figures of authority are usually present in reality, and the subject may be observed consistently to rely upon the means he has learned for dealing with them.

To make the best case for the incorporation into the ego system of figures of identification, we should, perhaps, consider those instances in which the conditions of identification continue for extended periods of time. The concentration camp offers a good example. Bettelheim (4) observed among prisoners some striking instances of identification with the guards, usually after the man had been an inmate for a long time; and Bettelheim, very properly it seems, referred to these phenomena as changes in the personality. The question would be, were these changes maintained after the prisoners were released? I do not know of any studies in which men observed in concentration camps were followed up after release. It seems reasonable to suppose that, just as there were wide individual differences in the tendency to identify with the guards, there would be individual differences in the tendency to persist in the identification after release. Among the factors favoring persistence, first importance might well attach to the presence in the personality of structures deriving from earlier introjection—structures to which the new object became assimilated. But if this were indeed the case, we would deal no longer with identification, but with introjection.

Probably one of the main sources of the widely held contemporary view that identification with the feared authority is an important source of the superego is Anna Freud's writing on "identification with the aggressor" (5). Here it is stated that this kind of identification

contributes to superego development, that it represents a stage in the development of the superego. Yet, her essay may actually be used to support the present view. For one thing, in those episodes which most clearly exemplify identification with the aggressive or dangerous object, i.e., episodes in which the child's identification with the external object is not complicated by the subject's own instinctual needs, Anna Freud says "it's not clear what became of the threat with which they identified themselves" (p. 122). She makes quite a leap from these episodes from child-analytic practice to the proposition that repetitions of the aggression or danger finally result in the setting up of the threatening agency "inside" the personality. She offers no evidence in favor of this proposition. She does, however, make quite vivid the view that the child "does not wholeheartedly accept this institution," and whereas there seems to be the assumption that later it will, she notes "it is possible that a number of people remain at this stage." In view of our present knowledge of the authoritarian personality, it might be said that a great many people remain at this stage of development; indeed one might question the view that this state of affairs is a stage at all, in the sense that it has a place in a progression toward the full acceptance by the individual of the critical agency. If individuals in whom "identification with the aggressor" is pronounced go on to develop a mature conscience, this is very probably due to the intervention of quite different determinants and not to an extension of this mechanism. In other words, this development would seem to occur in spite of the identification with the aggressor rather than because of it.

Fromm (9) has seized upon such accounts of superego formation as that given by Anna Freud in her chapter on identification with the aggressor to make

his point that the Freudian superego is an authoritarian one. If this account were all we had to go on, Fromm's point would, I think, have to be admitted. In Freud's own writings on superego formation (6, 7, 8), however, various other processes are given place; most central, it seems, is the introjection of parental figures as compensation for their loss as love objects. I would suggest that in the contemporary authoritarian personality there is less of the former punitive parent "inside" the personality than either Fromm or Anna Freud seems to imply. But this does not mean that such a personality is empty. There is actually a primitive superego deriving from earlier introjection; it is frequently to escape the onslaughts of this agency that authoritarian personalities seek to align themselves, through identification proper, with powerful external authorities.

There remains, however, the possibility that an identification initiated in one of our critical situations might become established in time as a means for controlling impulse. It does appear to be true that, in authoritarian personalities, the most extreme identification with authorities is accompanied by the most intense underlying hostility toward them; it is reasonable to ask whether one of the main functions of the former is not to inhibit or counteract the latter, and whether a device which thus becomes tied to an impulse does not thus gain the status of a fixture in the personality. Indeed, it may be asked whether such a fixture, originally designed to inhibit one impulse, might not be employed to inhibit others as well; in short, whether it does not become, in this way, a part of the superego. I do not wish to exclude such a possibility. But let it be noted that the conditions for this outcome are rather complex; they have to do both with the status of the personality before the crisis that

evokes identification, and with processes that operate after its occurrence. As suggested above, the new figure of identification may be assimilated by earlier introjections. In any case, the new figure is at least a datum of the individual's experience, something to be taken into account in the building of an ego system. But the emphasis here is upon this work of building rather than upon the initiation in a crisis of the unconscious copying reaction.

In sum, identification proper does not appear to be a very fruitful source of internalized structures in personality. To account for the superego and other internal agencies, we shall have to rely mainly upon introjection and other processes.

We should not take the too easy course of saying "introjection" whenever it appears that some feature of the environment has somehow found its way into the personality. It is unfortunate that Freud tended to do just this. He seems not to have cared much for Ferenczi's term "introjection"; he did not use it in any strict technical sense, but only in a general descriptive way as when he wrote "the external restrictions are introjected, so that the superego takes the place of the parental function" (8, p. 85). This kind of usage naturally has led psychologists to doubt that we need such a fancy word for a process that seems readily described in more familiar, generally accepted terms. R. W. White, in his influential *The Abnormal Personality* writes, "The concept of *introjection* has been considerably elaborated in psychoanalytic writings. The introjection of parental restrictions, however, is basically a learning process and may even be likened to simple conditioning. The child acts on some impulse, the behavior is punished (or punishment threatened), and this linkage with punishment causes the action to be internally inhibited thereafter" (19, p.

169). And White goes on to explain in similar terms how positive ideals may be inculcated. I think his is an accurate formulation of a very common phenomenon—a matter to which we shall return in a moment—but this phenomenon is not introjection, that is to say, it is not what we shall have to call introjection proper.

For a useful conception of introjection we should go back to the original formulation of Abraham (1), and to its vigorous following up by his student, Melanie Klein (10). Here we find the conception of a psychological taking in that is modelled after oral incorporation. When the object of love or imperious desire is withdrawn, or lost, or when its withdrawal or loss seems imminent, the subject may set it up imaginatively inside his personality, where he may, so to speak, have it for good. In infancy, when the boundaries between inside and outside are not yet clearly drawn and when the danger of losing objects upon which there is total dependence is subjectively acute, introjection is probably very common if not universal. And since the love objects of infancy and early childhood are major sources of frustration, they tend to be hated as well as loved; that is to say, they tend to be loved ambivalently. Frequently, it seems, the infant in striving to cope with his aggressive impulses, projects them onto the objects which are central in his scheme of things. Thus it is that the psychological meaning of the introjected objects may depend in considerable part upon what has previously been projected onto them. Although introjection is most characteristically a phenomenon of infancy, it may, its conditions being sufficient, occur at any time of life; in these instances the introjections are referred to, and assimilated by, earlier introjects of the same general significance. In my own experience—and I expect this is common—I have

never encountered a case in which introjection loomed large, either as a matter of historical fact or as a phenomenon to be dealt with in the transference, that did not present at the same time a picture of intense oral strivings and oral character traits.

Introjection as here conceived was called identification by Freud, most of the time. The following passage is typical, "If one has lost a love object or has had to give it up, one often compensates oneself by identifying oneself with it; one sets it up again inside one's ego, so that in this case object-choice regresses, as it were, to identification" (8, p. 86). And it is important to note that this is the conception that Freud relies upon in explaining how the super-ego becomes the heir of the Oedipus complex, i.e., identification with parents compensates for their loss as objects. It is true that Freud, most characteristically I think, referred to these phenomena as narcissistic—in contradistinction to hysterical—identification; and the distinction which he drew in this connection is most important in differentiating between introjection and identification proper. "The difference, however, between narcissistic and hysterical identification may be perceived in the object-cathexis, which in the first is relinquished, whereas in the latter it persists and exercises an influence, usually confined to certain isolated actions and innervations" (7, p. 160). This is the basis for the present view that in introjection the object "disappears inside" as it were, whereas in identification proper a continuing relationship with the external object is a distinctive feature. Perhaps it should be added that identification proper is intended to be a broader concept than hysterical identification.

Introjection, like identification, is an unconscious mechanism resorted to in a crisis. But the crisis is at once more

severe than, and qualitatively different from, that in which identification is evoked. The danger is perceived to be more extreme and the capacities for dealing with it are felt to be, as they actually are, far less adequate. For these reasons introjection occurs, most characteristically, in an earlier stage of development than does identification. In order for identification to take place there must exist, at least in rudimentary form, a conception of the self; this is not true of introjection.

The distinctive feature of the crisis in introjection is frustration in love; there is a loss or deprivation of love, or a real or imagined threat of such a loss. The object that is introjected is always one that has been loved ambivalently. In the oral and anal stages of development all object relations have this aspect of ambivalence; and it is probably true that all introjections either occur during these stages or have antecedents there.

Of course the kind of thing I have been trying to describe as introjection does not, so to speak, happen every day. We are talking about a mechanism that comes to the fore particularly in psychosis, and in severe psychosomatic disturbances. I favor making pretty categorical distinctions between this mechanism on the one hand and identification and normal character development on the other. But I do not favor any such categorical distinction between psychotic and normal people. Better to say that all of us, having been infants, having had oral and anal experiences that were in varying degrees shocking, have varying amounts of psychotic potential, and that it can still be rewarding to study the conditions under which this potential breaks into action.

Abraham made it clear that he considered a superego based on introjection as a "pathologically formed one." Freud, as indicated above, in describing the

superego in the *New Introductory Lectures* (8) gives the most central role to the self-same processes that Abraham makes so vivid in his study of the manic-depressive psychoses. And for his most vivid description of the superego at work, Freud goes back to the example of melancholia. Some reconciliation of Freud and Abraham may be achieved by recalling that for Freud the superego was, chiefly, one of those lesser evils that loomed so large in his scheme of things; primitive, blind, automatic, unconscious, it was to be got rid of, as soon as possible. This is not far from being something pathological, pathological at least in the sense of being maladaptive, undesirable, not healthy in any normative sense. The superego for Freud was not the same as conscience—not usually, at any rate; it was not the expression of our "higher natures." And Freud does not make the mistake of supposing that the processes by which the mature character is formed are the same as those which operate in superego formation.

It may seem that we have so restricted identification and introjection that we shall now be hard put to it to explain how so much of the culture, of the parental standards and prohibitions, *does* get firmly established in the personality. I think we need not worry. The processes described by White, above, cover after all an enormous amount of territory; for many psychologists they are sufficient to account for the whole of character development. Psychoanalysts interested in "developing an ego psychology" should realize that such a psychology has existed for quite a while. The bulk of psychological writing in the field of personality development may properly be put under this heading. Consider, for example, McDougall's (12) account of the development of the self-regarding sentiment, an account of conscience that has still to be surpassed, or

Murphy's (14) account of the development of the self, or Newcomb's (15) learning to take social roles. These writers are describing actual processes that are central to the normal psychology of character development. The fact that they do not accent unconscious processes, or distinctions between conscious and unconscious, that their views were little influenced by psychoanalytic theory, does not for a moment invalidate what they have to say about the particular ego processes upon which they focus. I am not suggesting that our knowledge of normal character development, of constructive interpersonal relations, and the like is complete, but only that we can restrict identification and introjection in the way that I have indicated without handicapping ourselves in the continuing search.

Now it is possible to see why I suggested that Mowrer's "defensive identification" was actually developmental and not an instance of identification. Accepting parental standards as a means for pleasing them and as a means for controlling impulse does not require the operation of any peculiarly psychoanalytic mechanism. The acceptance is by and into the ego system; the whole process may be largely conscious; standards accepted in this way may contribute heavily to the more or less mature conscience.

I propose, in conclusion, that we do two things. First, we should give more attention to those superego elements that have been based upon introjection. Second, we should consider that normal character development can be largely explained, without benefit of either identification or introjection, on the basis of common forms of learning. A child learns which of his actions please, and which displease, his parents, which win him love, which disapproval; he learns what reactions are effective in inhibiting those impulses which if allowed free rein

would lead to catastrophe; he learns how to regard himself from the way others regard him; and in building his ego system and his self-conception he learns what to keep and what to discard.

This conclusion presents me with an intriguing idea. Perhaps there are some of us psychologists who have not yet given allegiance to (I started to say "become identified with") any of the Freudian, anti-Freudian, or neo-Freudian movements, who still have need to come to terms with Freud and who would rather "be like him" than "to have him." Up to now, there have been two major neo-Freudian movements, the one based on the belief that Freud went too deep, the other on the conviction that he did not go deep enough. One possibility remains, that of going deeper and less deep at the same time. The argument of the new movement would be that Freud was not extreme enough; he stuck too close to the middle of the road. It would not be necessary to say which side of the road one would prefer. I am beginning to believe I mean this seriously. At least, I believe that if we can make clear and significant distinctions, and learn when to apply which explanatory concept, there is no reason why we cannot continue to make progress in ego psychology even while we continue the exploration of the unconscious.

REFERENCES

1. ABRAHAM, K. *Selected papers*. London: Hogarth, 1927.
2. ALEXANDER, F. *The medical value of psychoanalysis*. New York: Norton, 1932.
3. BALINT, M. On genital love. *Int. J. Psycho-Anal.*, 1948, 29, 34-40.
4. BETTELHEIM, B. Individual and mass behavior in extreme situations. *J. abnorm. soc. Psychol.*, 1943, 38, 417-452.
5. FREUD, ANNA. *The ego and the mechanisms of defense*. New York: International Universities Press, 1946.

6. FREUD, S. On narcissism. In *Collected papers*. Vol. IV. London: Hogarth, 1934. Pp. 30-59.
7. FREUD, S. Mourning and melancholia. In *Collected papers*. Vol. IV. London: Hogarth, 1934. Pp. 152-170.
8. FREUD, S. *New introductory lectures on psychoanalysis*. (3rd Ed.) London: Hogarth, 1946.
9. FROMM, E. *Man for himself*. New York: Rinehart, 1947.
10. KLEIN, MELANIE. Contribution to the psychogenesis of the manic-depressive states. In *Contributions to psychoanalysis 1921-1945*. London: Hogarth, 1948.
11. KNIGHT, R. P. Introjection, projection and identification. *Psychoanal. Quart.*, 1940, 9, 334-341.
12. McDUGALL, W. *Introduction to social psychology*. London: Methuen, 1908.
13. MOWRE, O. H. *Psychotherapy: theory and research*. New York: Ronald, 1953.
14. MURPHY, G. *Personality*. New York: Harper, 1947.
15. NEWCOMB, T. M. *Social psychology*. New York: Dryden, 1950.
16. REIDER, N. Psychodynamics of authority with relation to some psychiatric problems in officers. *Bull. Menninger Clin.*, 1944, 8, 55-58.
17. SYMONDS, P. M. *Dynamic psychology*. New York: Appleton-Century-Crofts, 1949.
18. TOLMAN, E. C. Identification and the post-war world. *J. abnorm. soc. Psychol.*, 1943, 38, 141-148.
19. WHITE, R. W. *The abnormal personality*. New York: Ronald, 1944.

(Received February 16, 1954)

A STATISTICAL MODEL FOR THE PROCESS OF VISUAL RECOGNITION¹

ARNOLD BINDER

Indiana University

Experiments dealing with recognition and similar perceptual phenomena have a number of characteristics in common, regardless of their theoretical antecedents. These common ingredients include: (a) a stimulating situation, (b) a set of instructions, (c) a general class of potential stimuli (out of which a particular one is chosen), (d) a response of the subject, and (e) a general class of acceptable responses (out of which one is made). The purpose of this paper is to present a statistical model of the recognition aspect of visual perception which is based on these common characteristics.

Before discussing the model it is of some advantage to define the terms *recognition* and *misperception* as they will be used throughout the treatment.

When an individual is shown a stimulus object and asked to name it, his response may be considered as either correct or incorrect in accordance with standards established beforehand. Accordingly, a response is correct if it corresponds to one of a set of specified class names (the set may consist of only a single class name), and incorrect if it does not so correspond. We will say that a subject has recognized a visually presented object if he responds with an appropriate class name of the object.

A generally useful way of establishing the class names which are to be considered correct in any given experiment is by statistical analysis of the frequency of different responses to various

presented objects. In this way the role of class names may be assigned to nouns which reach a critical response frequency. If, for example, 99.99 per cent of the people who look at a small, spherical, orange-colored object with an uneven texture call it an "orange" we can say that the appropriate class name is "orange" and define a correct response accordingly. For many purposes this method of establishing correct and incorrect responses is unsatisfactory either because of insufficiently high response frequencies or because of the desirability of defining certain low frequency responses as correct. As an example of the latter, the response frequencies to the stimulus object orange may be "orange" 99.99 per cent, "elephant" .001 per cent, "tennis ball" .003 per cent, "watermelon" .001 per cent, "fruit" .003 per cent, and "food" .002 per cent. Instead of merely specifying "orange" as the correct response on the basis of its very high frequency of occurrence, we could present the stimulus object to a new set of subjects and ask them, successively, "Could this be an orange?" "Could this be an elephant?" "Could this be a tennis ball?" and so forth until every response is substituted in the phrase. By this method correct responses could be defined on the basis of the response "yes" to the question, "Could this be a _____?" reaching a high frequency among the subjects. (Note that both these methods have been used for the establishing of *F+* responses, *F-* responses, populars, and originals on the Rorschach although the rationale has not always been explicitly stated; see 1, 11, 12.)

A subject who is confronted with a

¹ I am grateful to Drs. A. M. Buchwald, C. J. Burke, D. M. Ehrman, and D. G. Ellson for critically reading the manuscript and offering suggestions toward its improvement.

stimulus object under instructions to name it may react in one of four mutually exclusive ways. He may respond with one of the appropriate class names, respond with a class name which is not correct, respond in a manner irrelevant to the task (such as describing certain Civil War battles), or he may make no verbal response. We say that the individual has misperceived the object only if the failure to recognize consists of an erroneous class name. Thus, if the individual calls the above object an "orange" we say that he has recognized it, but if he calls it an "elephant" (or even an "apple" or a "tennis ball") we say that he has misperceived it. If the subject responds in a way other than naming the object or gives no response, we can make no statements as to misperception since the conditions of the instructions have been violated.

THE MODEL

Let us start with a primitive set, the elements of which we shall call stimulus objects; this set corresponds to the general class of potential stimuli mentioned in the introduction. We assume that each object in this general or primitive class has a denumerable set of characteristics or attributes by which it is identified. On the basis of these characteristics or attributes we can divide the general class of objects into various subclasses in such a way that all objects in any one subclass have one or more attributes in common. (If two or more such subclasses have exactly the same set of common attributes we will use only the one subclass among these of greatest element cardinal number.) Once this is accomplished, it is a simple matter to give class names to preferred subclasses.

For terminological clarity we shall henceforth reserve the term "class" for referring to a set of objects which are grouped together on the basis of a com-

mon array of attributes, and use the term "collection" to refer to the set of all classes under consideration. The symbol " V " will be used to designate the collection of classes and the symbol " K_i " the set of objects falling in the i^{th} class (i going from 1 to the total number N of classes in the collection, assuming of course that they may be arranged in a sequence). Where the classes have been assigned class names, the member objects of the classes, as well as the classes themselves, may be referred to by these names.

Thus, if our collection consists of the classes designated by the class names fish, book, woman, pencil, orange, pear, ball, cherry, chair, pea, planet, grapefruit, every object which contains the set of attributes spherical, about three inches in diameter, orange colored, rough textured, and green stemmed belongs to the K_1 class designated by the name "orange." Different classes may have a few or even many defining attributes in common, although they must differ in regard to at least one attribute. An attribute may of course be common to the members of only one class, the members of a few classes, the members of many classes, or even the members of the entire collection of classes.

Assuming that the objects coming under each of the class names in the collection above possess their usual attributes, it is apparent that no class (such as fish) includes another class (such as orange) as a subclass. But this is obviously not a necessary restriction for the collections with which we may deal. For example, our collection could contain the classes orange (say K_1), citrus fruit (say K_2), food (say K_3), navel orange (say K_4), and "juice" orange (say K_5), so that

$$K_1 \subset K_2 \subset K_3$$

and

$$K_1 \supset K_4$$

$$K_1 \supset K_5;$$

assuming, of course, a usual array of attributes.

Since an object is a member of a class if, and only if, it possesses the set of attributes characteristic of the class, we can define the order of a K_i class in a collection by the inverse of the ratio of the number of attributes common to all member objects in the class to the total number of attributes in the system (provided that the total number of attributes is finite). The members of the class food have fewer attributes in common than the members of the class citrus fruit, since the members of the latter class possess every attribute of the members of the former class plus a few more specific attributes. Thus, in a given collection, the class food is of higher order than the class citrus fruit. An object containing, say, ten attributes may be a member of a lower order class containing, say, nine common attributes and at the same time a member of a higher order class containing, say, three common attributes.

As the sequence goes from an attribute common to the members of only one K_i class to an attribute common to the members of all classes in V , an attribute moves from one with greatest differentiating power to an attribute with minimum differentiating power. The differentiating power of an attribute is inversely proportional to the degree of ambiguity involved in assigning the object possessing the attribute to a particular class. In the collection presented above containing fish, book, woman, pencil, orange, etc. the attribute of "edible" has relatively little differentiating power and the attribute "wears dresses" has maximum differentiating power.

The task involved in this model is the assignment of an object, selected from the primitive set, to an appropriate class (K_i) on the basis of a knowledge of all or some of the object's attributes. Assignment must be made to

a class whose members have *all* of the known attributes in common. Regardless of the quantity or nature of its total set of characteristic attributes, a class may be selected for assignment only if it has at least the known or available attributes as a part of its defining array. The number of possible assignments is of course identical to the number of classes in V whose member objects have in common all the available attributes. The degree of accuracy with which this assignment or designation may be done depends principally upon the differentiating power of the attributes which are available. In the collection of classes containing fish, book, woman, pencil, orange, etc. (designated above), for example, there would be considerable error in assigning an object to a class if it were only known that the object was spherical. There would be much less error if it were known that the object was spherical and about three inches in diameter; and no error if it were known that the object was spherical, about three inches in diameter, orange colored, and rough textured.

One last point. The model is applicable to those situations in which each class has a different probability of occurring in the form of a specific member object as well as to those situations in which these probabilities are equal.

METHODOLOGY²

The above model is well suited for quantification by the methods of infor-

² This approach to mathematical analysis assumes that, where there is a choice, an object will be assigned to the lowest order class in V of which it is a member. A necessary concomitant of this is that two items of information about attributes may, in various recognition situations, serve as attributes in differentiating classes for assignment purposes. These items are: "All of the potential attributes of the given object are available," and "All of the potential attributes of the given object are not available."

mation theory (16). Information is the same as the amount of uncertainty or the freedom of choice in assigning an object to a class. Uncertainty depends not only on the number of possible alternative choices (or placements) in the source after a given message set (or set of attributes) is known, but also on the probability associated with each alternative in a given context. For purposes of illustration let us consider the process as reflected in a noiseless system.

Information or uncertainty is measured by entropy defined as

$$H = - \sum P_i \log_2 P_i,$$

where P_i is the probability of occurrence of an object from the i^{th} class, and i ranges from 1 to the total number (R) of K_i classes whose member objects possess all of the particular attributes which are available (plus, in most cases, others which are not known). (Note that the range of i may just as well be considered 1 to the total number (N) of classes in V , provided the probability of zero is assigned to those classes whose members do not possess the set of available attributes.) The units of entropy are bits of information. It is evident that entropy is a measure not of the information which is available regarding the identity of the class of a particular object, but of the average amount of information still necessary before the object can be uniquely assigned to a class. In the recognition model each attribute is a message and the entropy varies with the set of known attributes. Suppose we have the collection of classes designated by the class names ink, pen, deer, paper, dog, cat, tumbler, human, goldfish, duck, fan, book, rabbit, brick, bush, hammer, light bulb, worm, tiger, turtle. If an object from one of these classes is concealed in a room and if it is only known that this inaccessible object moves spontaneously, there is a good deal of uncertainty as to its class. Or, in other

words, much more information is needed to uniquely identify the class of the object. The entropy would reflect both the fact that there are a number of possible classes which meet this specification and perhaps that objects from certain classes are more likely of occurrence in the room than others. If it is also known that the object moves about on four legs, the number of possible alternatives and thus the entropy decreases. Suppose that another attribute specifies the object in question as being a pet of Mr. M. whose pet aggregation consists of cats, dogs, and rabbits. Suppose further that Mr. M. has an equal number of dogs and rabbits, but twice as many cats as rabbits; and that each individual animal has an equal probability of being selected for concealment purposes. The probability that the animal is a dog when all this information is known is $1/4$; that it is a cat, $1/2$; that it is a rabbit, $1/4$. The entropy of the source would, therefore, be

$$H = - (1/4 \log 1/4 + 1/2 \log 1/2 + 1/4 \log 1/4) = 3/2 \text{ bits.}$$

Whenever the probabilities of occurrence of the various classes are equal (i.e., all P_i are equal) the formula for entropy reduces to

$$H = \log_2 R,$$

where R is equal to the total number of classes (or alternatives) which have the set of available attributes. If an object is assigned to a class when a class of which it is a member is not uniquely specified by the known attributes, there is a probability associated with the correctness of placement which is always less than one. In the case where each class has an equal likelihood of occurrence, the probability of correctness of placement is $1/R$. The matter is considerably more complex when the P_i 's are not equal, in which case the probability of each possible assignment of class must be taken into consideration.

Consider the example above regarding Mr. M.'s pets. Suppose that in a similar situation repeated 100 times, the individual assigning the objects to classes always responded "dog" after the above series of attributes were given. We can thus state that the probability of the same response being made in the same situation is very close to one (we can, obviously, never be completely sure). The probability of being correct can easily be shown by conditional probability to be equal to

$$(1)(\frac{1}{4}) + (0)(\frac{1}{4}) + (0)(\frac{1}{4}) = \frac{1}{4}.$$

Now let us assume that the individual uses a frequency of designation of a particular class name equal to $P_i n$ where P_i is the objectively measured probability of occurrence of the given class and n is the number of repetitions of the set of conditions. The probability of his being correct in this case would be approximately equal to (approximately because the response probabilities cannot be determined precisely, but only estimated from previous performances)

$$(\frac{1}{4})(\frac{1}{4}) + (\frac{1}{2})(\frac{1}{2}) + (\frac{1}{4})(\frac{1}{4}) = \frac{3}{8}.$$

The reciprocal of the antilogarithm of the entropy is a very close approximation of the probability of correctness whenever the individual's probability of each designation is equal to the objective probability of occurrence of an object from the class in question. It is exactly equal to the probability of correctness whenever each class has an equal probability of occurrence, regardless of the individual's probability of making each assignment. The former statement is readily established by a comparison of the expansions by Taylor's Theorem of the Napierian logs of

and the latter by a substitution of the value $1/R$ for each P_i in the two expressions

$$\sum_{i=1}^R X_i P_i \quad \text{and} \quad \prod_{i=1}^R P_i P_i$$

(where X_i is the probability of the individual's assigning an object to the i^{th} class on the basis of the available attributes and $P_i = 1/R$ is the probability of occurrence of the i^{th} class out of the R possibilities).

For many purposes it is not necessary to know the probability of accuracy of assignment since the objective measure of entropy is proportional to this probability (assuming that the pattern of assignment probabilities does not vary) and can be used comparatively to evaluate differences in the probability of accuracy of placement.

THE PROCESS OF RECOGNITION

For its specific reference to the phenomena of perception it is convenient to refer (with one exception) to the attributes or characteristics of objects as cues. The exception is in those attributes which belong to all of the classes (K_i) in V ; these will not be referred to as cues. Thus, cues may be defined as elements of any variety which differentiate the members of certain classes from the members of other classes.

Assignment of an object to a class is accomplished by a response consisting of a class name. The model implies for recognition, then, that a person, on observing an object under instructions to name the object when he recognizes it, responds with a class name after he has enough cues available to identify the object. The use of the data in the form of cues functions to lower the observer's uncertainty as to the identity (accepted class name or names) of the object. The evidence available indi-

$$\sum_{i=1}^R P_i^2 \quad \text{and} \quad \prod_{i=1}^R P_i P_i,$$

cates that when the number of cues is large, subjects accumulate a portion of these cues successively. Although it is obviously not a necessary restriction, we assume, for expository as well as other purposes, that subjects accumulate all cues in a successive manner, so that each cue reception consumes a finite interval of time. In the case of the orange, the sequence might go something like: spherical, about three inches in diameter, orange colored, rough textured, "Aha, an 'orange.'" This is perhaps an oversimplification since the subject does not always have the same level of confidence in the correctness of his response. He is more or less certain of the correctness of his response; we will call this his subjective probability of accuracy of recognition. After the sequence of cues representing the orange, therefore, his response might more accurately be represented by the sentence, "Aha, it is almost certainly similar to the objects called 'oranges.'"

Accepting the successive nature of the accumulation of cues, we can equate the more and more complete specification of a class by the series of cues with the decreasing number of possible alternatives. Perhaps there are 1000 possible classes whose objects are spherical; 100 whose objects are spherical and orange colored; 10 whose objects are spherical and orange colored and about three inches in diameter; and only one whose objects are spherical, orange colored, about three inches in diameter, and of rough texture. The uncertainty at any particular choice point, of course, is a function of the probability of occurrence of each alternative as well as the number of possible alternatives.

No matter what sort of a cue system is used there is likely to be in presentations of stimulus objects a redundancy of cues. Even though the orange is recognized after the four cues above are available, there are a few additional cues such as the existence of a green

stem and perhaps a navel which might have been used for differentiation had the object not been identified.

It would seem desirable to emphasize the distinction between uncertainty, as used here, and the concept of subjective probability. Subjective probability may be defined as an individual's personal estimate as to the probable accuracy of a given perception. This would seem related to uncertainty as objectively measured, which is of course a function of both the number of alternatives and the probability of occurrence of each class. In a given individual, one would expect subjective probability to be inversely proportional to objectively measured uncertainty, although the same amount of objective uncertainty does not necessarily lead to the same degree of subjective probability of accuracy in different people. Some, for example, may be quite certain as to the accuracy of perception after a certain three cues, while others in the same context of conditions are equally certain only after these same three cues plus two others. In those situations in which an individual is asked to name an object whenever he "recognizes" it, response will, presumably, occur at the point that a given level of subjective probability is reached in the cue gathering process. This level will vary somewhat from situation to situation, depending possibly upon the consequences of the response.

According to our assumption (of convenience) the time spent on each cue is extremely small, but nevertheless finite. Of course the cue-time sequence is a function of what is considered a cue. Suppose we say that a person, on viewing a flat surface, can obtain at the most one cue from each one-sixteenth of a square inch area. Let us say, further, that the time of delay at each such cue accumulating area is one-hundredth of a second, so that his center of attention shifts only after that length of time.

Under these conditions one would get the greatest degree of differentiating power in terms of analysis by calling the effect of each one-sixteenth square inch area a cue. But this is not crucial. One could just as well designate each square inch of area a cue, so that the time of delay would be sixteen-hundredths of a second. The fact that we have one cue now in the place of the sixteen which we had previously is only of methodological importance in terms of the particular objects which are to be differentiated. The only requirement for an element to be designated a cue is that it differentiates certain objects from certain others, whether the differentiating process is fine or coarse. The extension of the notion of cue to such features as form, color, and texture amounts to nothing more than an extension of the above reasoning. Although a finer breakdown of cue elements would undoubtedly be desirable for many experimental procedures, it seems likely that this would do little more than lead to cue redundancy for most objects commonly encountered.

RELATIONSHIP TO PREVIOUS WORK IN PERCEPTION

Let us first consider the relationship between the definitions of recognition and misperception presented here and the discussions of other writers on related matters. A phenomenon that may be called "a simple visual experience" is often said to occur whenever an individual is confronted with light of simultaneously varying intensity and wave length, and he can only respond by specifying the various psychological counterparts of the physical energy (brightness, hue, saturation, etc.). In Gibson's (9) terms, he can respond by describing the "stimulus correlates of the retinal image." A visual experience is generally thought of as becoming meaningful to an individual when he can not only describe these psychophysi-

cal characteristics but can relate the total stimulus configuration to previously experienced phenomena. He can thus speak, for example, of the familiar characteristics of an object, its general uses, or perhaps its emotional significance. Recognition takes place when a visual experience becomes meaningful.

It is rather interesting that subjects in recognition experiments are typically not asked to indicate recognition by specifying these various associations. An object's meaningfulness is usually specified by an individual's responding with a class name (or making some similar designatory response) when confronted with the visual stimuli (see 2, 6, 7, 14, 15). Implicit in the reasoning behind this substitution is the assumption that when an individual has responded with an accepted class name of the object he could also have given other and meaningful associations. (The accepted class names defining a correct response are most often established in the statistical way described above although this is very seldom made explicit.) In the case of the object, orange, this implies that a person responding "orange" has not only learned the connection between object and name in a rote manner (similar to the association of two nonsense syllables), but could further specify that the object has juice inside, that it is eaten, that it grows on trees, that it has a sour-sweet taste, and so forth.

This manner of approach and the one described earlier in the paper lead to the same kind of experimental methodology. The one in this paper, however, would seem to have advantages in terms of freedom from loosely defined terms and concepts.

Now we consider some of the experimental evidence and the theoretical formulations of a number of investigators which are related to other aspects of this model.

Buswell, on investigating the eye movements concomitant with the per-

ception of pictures, found that two perceptual patterns were evident in observations of the pictures by his subjects; the one consisting of relatively large eye movements with short pauses and the other consisting of smaller movements with longer periods of fixation. The former seemed to function as a general survey of the principal features of the picture and the latter as a close inspection of important pictorial sections. He concluded that the "short fixations are apparently related to the simpler processes of visual perception, whereas the longer fixations seem to indicate a mental process of reflection or at least a higher degree of interest than is occasioned by the ordinary survey of the picture" (5, pp. 142-143). The eye movements of the general survey variety with the shifting points of fixation, then, would seem to serve as sequential information gathering facilitators in the sense of making available new sets of cues when all of the cue data from a previously focused upon area have been exhausted.

There is evidence to indicate that these movements and fixation points are not randomly distributed, but influenced by the set which the subject has prior to the observation and this set as modified by the accumulation of informational cues. In Buswell's study, for example, changes in the preliminary directions markedly altered the patterning of eye movements, the manner of approach, and the duration of the eye fixations. The revision of sets as additional cues are received is nothing but a logical extension of this process since the point at which preparation ends and perception begins is quite arbitrary. Bruner has called these sets or tendencies to attend selectively to certain aspects of the environment rather than others "expectancy hypotheses," and states that they shift "... in a direction partly determined by internal or personological or experiential factors

and partly on the basis of feedback from the learning which occurred in the immediately preceding, partly unsuccessful information-checking cycle" (4, p. 124).

In examining a complex object one person may name it after receiving n_1 cues while another names the object only after receiving n_2 cues (where $n_2 > n_1$). Assuming, first, that the two individuals receive the same sequence of cues and have the same background of relevant experience and, second, that the total number of cues necessary for unique specification is n_3 ($n_3 > n_2$), it is obvious that over a period of time the former individual is more likely to misperceive or misrecognize similar objects than the latter because of the entropy differences (provided they do not alter their patterning of responses, of course).

The hypothesis that different people respond at stages of cue reception differing in level of objective uncertainty can be deduced from the formulations presented above. This is essentially identical to Hilgard's (10) hypothesis of the tendency for achieving clarity when the cues do not fully define the objective situation. He emphasizes that people tend to use few cues in making anticipatory guesses as to the stimulus object, even though forthcoming cues might completely and uniquely specify it. Miller (13), in referring to the tendency to jump to perceptual conclusions before all cues are received, goes even further in stating, "This process of inductive belief—jumping to conclusions from inadequate premises—is the most essential cognitive process in the organism" (p. 278).

APPLICATIONS OF THE ANALYSIS

Ambiguous stimuli are very widely used at the present time for experimental purposes in both the laboratory and the clinic. Let us examine what implications this analysis has for the term

"ambiguity" as it is used to describe such stimuli.

We may consider ambiguity as it is reflected in a single cue and as it is reflected in the entire constellation of cues available from an object. Any ambiguity attached to the identity of a single cue results from noise in the system. That is, as a result of certain idiosyncracies of transmission (external or internal) random errors (noise) are introduced into the system so that when a cue is received there remains uncertainty as to the true nature of the cue. (By "true" we mean of course the cue corresponding to the stimulus conditions defined by consensual agreement.) In information terms this is the uncertainty in the message after the signal is received and is called "equivocation." Equivocation would seem to be a factor of importance in the recognition model only when a subject has optical or brain disturbance or where the situational illumination is less than satisfactory, but of little consequence in most circumstances.

Ambiguity, as the term is used to refer to certain aspects of projective tests and tachistoscopic presentations, depends upon the acceptance of the condition whereby fewer cues are available than are needed for the unique specification of a class name. The condition depends upon the definition of a person other than the observer or upon an interpretation of the observer. This definition or interpretation is necessary for ambiguity; otherwise the subject would presumably use the lesser number of cues in specifying a class name in V whose member objects are larger in number and have less attributes in common. In the terms defined previously the subject is forced to use a class lower in order than one that is maximally differentiated by the cues. (Another way of looking at it is that certain higher order classes are ruled out of V by this process of ambiguity

specification.) Suppose one only had the cue "spherical," for example. If the situation is considered as one containing less than the full number of cues necessary for unique specification in a collection V composed of classes K , whose member objects have many more attributes in common, the object could belong to such classes as ball, orange, fortune-teller's crystal, or even planet. The uncertainty involved in the choice indicates the level of ambiguity. If, on the other hand, the situation is not considered as consisting of reduced cues, the object may be specified only by some such term as "sphere." In the case of the Rorschach, a subject could say "inkblot" to each of the ten cards if the experimenter had not defined the situation as one of reduced cues.

We may think of an ambiguous stimulus condition, then, as one in which the uncertainty cannot be reduced to zero. This is somewhat artificial since a person must do better (in terms of more specific designation) than the available cues would warrant.

Ambiguity or lack of "structuring" as it is more commonly referred to in projective practice is quantifiable by the techniques described in this paper. Of course it can be done only in those cases where one can specify (or approximate) the values of the various P_i 's.

From the viewpoint of the individual who responds before all the cues necessary for unique designation are received, every object is in a sense ambiguous. That is, uncertainty is never equal to zero. But this is certainly not the sense in which the term ambiguity has historically been used. This formulation depends upon characteristics of a certain set of observers and does not differentiate kinds of stimulus conditions.

Different subjects, on the basis of their usual methods of operation, their preliminary expectations, and their interpretation of the test instructions, respond in ambiguous situations at dif-

ferent levels of uncertainty. There are subjects who will respond verbally as a sign of recognition even though a very large number of alternative concepts are possible at the given stage of cue specification; similarly, there are subjects who will respond only when the concept is all but completely specified. In any particular experimental situation it is an easy matter to establish the average uncertainty at the point of response of the subjects. It seems likely that most normal people fall within a restricted range (under most conditions) in terms of the uncertainty they will "tolerate" in responding. On the Rorschach, as an example, this standard is roughly described by the fact that people tend to name an object on the basis of a single cue only when that cue is form. However, it does happen that a combination of other cues reduces the uncertainty of the source about the same amount as one form cue (provided the form cues are indefinite rather than contradictory), and people even with this minimum standard can achieve concepts on the basis of the specification by these combined cues. Thus, they might call Card X an "underwater scene" or Card IV an "animal skin."

People who "tolerate" little uncertainty may respond only after a series of cues all but uniquely specifies the object. Their behavior may indicate, in fact, that they do not accept the instructions relative to reduced cues. These people can deal with the Rorschach blots as only slightly more unstructured than photographs. In this sense the model provides a definition of "stimulus bound."

The application of the statistical methods of information theory to the process of recognition facilitates the formulation and/or testing of various hypotheses. Among these are: (a) In tachistoscopic presentations, objects which are completely specified (zero uncertainty) by very few cues (such

as an alligator) will be misperceived less often than objects which are only specified uniquely by a larger number of cues (such as a deer), even though both sets of objects are of equivalent complexity. (An indirect confirmation of this came in a recently reported experiment by Eriksen [8] who found that objects could be located in an object field of given attribute heterogeneity faster when more attributes uniquely differentiated the target object from field objects.) (b) Schizophrenics misperceive more often than normals because of their disinterest in and inattentiveness toward the informational cues coming from the objects of reality. (c) Overcontrollers (as defined by Block and Block, 3) perceive better (make fewer errors of recognition) as a result of a tendency to react only when situations are well defined. In order to test such hypotheses as the latter two, it is first necessary to define a collection (the possible alternatives prior to any cue transmission) and then isolate out a series of cues which will progressively delimit the number of possible alternatives. The collection must be specified beforehand in order to make the calculations of the various possible alternatives and their associated probabilities of occurrence.

In an experimental test of the hypothesis that Group A perceives better than Group B for one of the above reasons, for example, we would present the cues one by one to the various subjects in order to determine at what point they respond. Suppose we have a population of 128 objects all with equal probability of occurrence, and suppose that every cue represents one bit of information. The subjects are free to name the object after having received enough cues to satisfy them as to its identity. The entropy of the source may then be determined at the choice point so as to serve as a measure of the likelihood of misperception in one

of the forms discussed previously. Thus, if one subject names the objects after three cues are presented, the entropy of the source would be

$$-\sum P_i \log_2 P_i = \log_2 R \text{ (since all } P_i \text{ are equal)} = \log_2 16 = 4 \text{ bits}$$

at this choice point. While if he names the objects only after seven cues are presented, the entropy of the source would be 0 bits. The 4 and 0 are, of course, measures of the uncertainty of the identity of the stimulus at the time at which response is made. We cannot say that a subject is right or wrong in his response to a particular object if he has chosen one of the alternatives which meet the required specifications, but we can state the probability of his misperceiving objects of the given complexity if he responds after a certain number of cues.

SUMMARY

A statistical model of the process of visual recognition has been presented. In the model, objects are assigned to classes on the basis of their attributes and classes are defined in terms of the common attributes possessed by their member objects. Objects in the same class have exactly the same set of attributes, while objects in different classes differ in regard to at least one attribute. The task involved in the model is the assignment of an object to a class on the basis of a known set of attributes. The available attributes determine the number of possible alternative classes to which the object may belong and, concomitantly, the statistical uncertainty of the object's class name. A few of the simpler notions of Shannon's (16) information theory were introduced to facilitate methodological treatment.

REFERENCES

1. BECK, S. J. *Rorschach's test*. New York: Grune & Stratton, 1944.
2. BITTERMAN, M. E., & KNIFFEN, C. W. Manifest anxiety and "perceptual defense." *J. abnorm. soc. Psychol.*, 1953, 48, 248-252.
3. BLOCK, J., & BLOCK, JEANNE. An investigation of the relationship between intolerance of ambiguity and ethnocentrism. *J. Pers.*, 1950-51, 19, 303-311.
4. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951.
5. BUSWELL, G. T. *How people look at pictures*. Chicago: Univer. Chicago Press, 1935.
6. EAMES, T. H. A study of the speed of word recognition. *J. educ. Res.*, 1937, 31, 181-187.
7. EAMES, T. H. The speed of object recognition and of word recognition in groups of passing and failing pupils. *J. educ. Psychol.*, 1947, 38, 119-122.
8. ERIKSEN, C. W. Object location in a complex perceptual field. *J. exp. Psychol.*, 1953, 45, 126-132.
9. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
10. HILGARD, E. R. The role of learning in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951.
11. KLOPPER, B., & KELLEY, D. M. *The Rorschach technique*. Yonkers-on-Hudson, N. Y.: World Book Co., 1946.
12. McREYNOLDS, P. Perception of Rorschach concepts as related to personality deviations. *J. abnorm. soc. Psychol.*, 1951, 46, 131-141.
13. MILLER, J. G. Unconscious processes and perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951.
14. POSTMAN, L., BRONSON, W. C., & GROPPER, G. L. Is there a mechanism of perceptual defense? *J. abnorm. soc. Psychol.*, 1953, 48, 215-224.
15. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, 43, 142-154.
16. SHANNON, C., & WEAVER, W. *The mathematical theory of communication*. Urbana, Ill.: Univer. of Illinois Press, 1949.

(Received March 19, 1954)

THE INNSBRUCK STUDIES ON DISTORTED VISUAL FIELDS IN RELATION TO AN ORGANISMIC THEORY OF PERCEPTION¹

HEINZ WERNER AND SEYMOUR WAPNER

Clark University

More than half a century ago, Stratton (20), in his work on experimental inversion of the visual field, introduced a novel method into perception psychology; but in spite of its great potential importance for perceptual theory, little use was made of this method during the following decades. Even later on wherever Stratton's work was followed up (5, 18), the observations and the conclusions were centered mainly around the topic of sensorimotor adjustment, re-education, and so on. Only relatively recently has the "method of visual distortion" been utilized for the explicit purpose of understanding the genesis and the nature of perception as such.²

Basing his work on original studies by Erismann (4), Kohler of the Innsbruck Laboratory has published results of investigations (12, 13) which are concerned with stages of perceptual adaptation to a world distorted in various ways by mirrors or prisms; these studies appear to us to constitute a very promising experimental approach for the advancement of perceptual theory. Though one might have wished that the Innsbruck psychologists had applied greater rigor in their experimentation and technique of measurement, the work as it stands, in its richness of observations and relevant findings, must be looked

upon as of direct relevance in the controversy between sensorial and organismic theories of perception.

In order to demonstrate the significance of the Innsbruck studies in regard to this controversy, we shall first present some of the most important methodological aspects and findings of Kohler's work; then we shall point to certain difficulties in interpretation of these findings for a sensorial theory of perception but not for an organismic theory; and finally, we shall discuss how one particular organismic theory, viz., the sensory-tonic field theory of perception, deals with findings of this sort, and how, in turn, Kohler's work aids in the generalization of constructs derived from this theory.

METHODOLOGY AND IMPORTANT FINDINGS OF THE INNSBRUCK STUDIES

There are three significant features of Kohler's methodology which should be noted. One is the extended duration of several experiments. The duration of the first part of the experiment, during which distorted media (mirrors, prisms, etc.) were worn by a number of the subjects, varied from a few days to 124 days.³ A second feature is the utilization

¹ The experimental work conducted at Clark and referred to in this article was supported by research grant MH-348 from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

² Gibson's work (6, 7) is, at least theoretically, closely related to this problem area; one may note that in one of his studies on adaptation to curved lines, use was made of binocular prismatic wedges.

³ It should be noted that Ewert's three subjects wore inverted lenses for 14 to 16 days, Snyder and Pronko's one subject for 30 days. The emphasis in these studies was on sensorimotor performance. In this connection it should be pointed out that in some experiments Kohler reversed the visual scene but did not invert it as Stratton, Ewert, and Pronko did. In several articles (13) Kohler has presented a lucid picture concerning developmental stages of perception during the course of adjusting to such a left-right reversed visual environment.

tion of "half lenses" or "half prisms" instead of full-length distorting media; by this technique the visual field becomes divided into two areas, one distorted and the other undisturbed. The third feature is the particular emphasis of the experimenter on the study of aftereffects, i.e., on careful phenomenological descriptions and perceptual tests after the removal of the distorting media.

Kohler's pivotal problem is his unique concern with the question of whether or not perceptual adaptation can occur in a differentiated manner, i.e., whether stimuli are perceived differentially depending on the introduction of several (rather than one) ocular-postural frames of reference.

To exemplify Kohler's method and findings we may discuss one of his experiments concerned with the effect of ocular-postural frames of reference on space perception, and another, concerned with the effect of ocular-postural frames of reference on color perception.

The following is a representative experiment on space perception. For 50 days an observer wore spheric half prisms attached to a frame in such a way that by looking upwards the subject gazed through the distorting prism whereas by looking downward there

was no prism and thus the gaze was free of distortion. There was clear evidence that the perceptual adaptation occurred differentially with respect to the two ocular-postural conditions: in time, with the gaze upward, the subject compensated increasingly more for distortion (shifts in location and changes in curvature); with the gaze downward, at first there were distortions in the opposite direction from those occurring under the upward gaze. These decreased considerably during the course of the experiment. Moreover, if, after a sufficient period of adaptation, the prisms were removed, the subject's visual world was found to be broken up into two perceptual views depending on whether the subject's gaze was upward or downward: upward gaze, compared with downward gaze, distorted the objects considerably more (in reverse direction to the distortion by the prism). In other words, the same retinal stimulation yielded different percepts depending on the one or the other ocular position.

The adaptation to curved lines, right angles, etc., and the aftereffects were measured by the so-called adjustment method: e.g., a test line was varied in curvature until it appeared straight to the subject; analogously, a rhombus

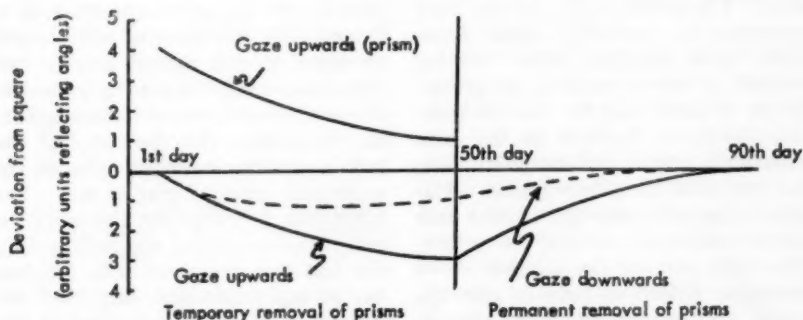


FIG. 1. Diagrammatic representation of effects and aftereffects of wearing half prisms on perception.

was systematically changed in its angles until it was judged a square. Figure 1 is a simplified diagram representing changes occurring in one subject during 50 days of the wearing of half prisms and of the aftereffects during the subsequent 40 days.

The abscissa represents the number of days; the ordinate indicates deviations in arbitrary units from standard, e.g., as depicted here, angular deviations from standard square. Thus, points on the ordinate above zero represent acuteness of the lower left angle of the test rhombus whereas points below zero represent obtuseness of that angle (Kohler constructed analogous diagrams for curvature of lines, etc.). The diagram contains two parts—left and right. The upper left curve refers to deviations occurring with the upward gaze (prism view); it shows decrease of deviation during 50 days. The two other curves at the left part represent measurements taken during short test periods under temporary removal of the prisms; the low curve refers to free gaze in the upward direction, the middle curve to free gaze downward. Inspection of the low curve (gaze upward under temporary removal of prism) shows an aftereffect due to previous wearing of the prisms. This aftereffect grows with the increase in days during which the prisms were worn. The middle curve (broken line) represents the deviations under downward gaze sampled after varying amounts of time of wearing the prisms. It can be noted that for the first three days deviations increase in the same manner for upward and downward gaze, but that after three days greater deviations occur only with the upward gaze under temporary removal of prisms. The right part of the diagram shows analogous differences between gaze upwards and downwards during the readaptation period, i.e., during the 40 days after permanent removal of the half prisms.

That different percepts occur with identical retinal stimulation but differing ocular postures is a conclusion which Kohler attempted to demonstrate not only in regard to perception of space but also of color. In his experiments on color the subjects wore, during a period varying from 20–60 days, glasses which were colored in the following manner: each eye piece contained a glass of which the left half was blue and the right half, yellow. Thus, the subject saw everything in blue color when looking to the left, and yellow when looking to the right. At the beginning of the experiment, Hering's afterimages were quite noticeable; they consisted in a short-lived strengthening of the blue hues at left after subject has looked through the yellow glass at right, and vice versa. These "orthodox" aftereffects decreased from day to day: the colors became increasingly paler. That is, contrary to what one should expect from the laws governing afterimages, a decrease of both complementary colors occurred within the foveal regions. Although it is to be regretted that quantitative data on the color adjustment are not given in Kohler's monograph, at least the general results of the measurements are stated. For these measurements a kind of color variator was used; the essential part of the apparatus consisted of a ground glass, the color of which could be varied from a neutral gray in two directions: toward increasing yellow on the one side and toward increasing blue on the other. The observer had the task of determining which color on the continuum appeared gray to him. The adaptation to the yellow glasses at the end of the experiment was indicated by the fact that, compared with the first day of experimentation, only "half the amount of blue" (as measured on the scale) was necessary to compensate for the yellow; the adaptation to the blue field, on the other hand, was almost

perfect. After removal of the glasses, stimulation of the fovea yielded two different color impressions depending on the direction of gaze: the equivalence point for gray under left gaze was a color well within the objectively blue range, and under right gaze, a color within the yellow range.

ORGANISMIC VERSUS SENSORIAL INTERPRETATION OF KOHLER'S WORK

The attempts to interpret Kohler's findings can be focussed either on the motor readjustment of the person while wearing prisms or on the perceptual changes occurring during that time.

As long as one emphasizes the first aspect, that of increasing adequacy of response to a disarrayed world—as most investigators have done—no particular theoretical difficulty arises, the problem being mainly to describe adequately the course of relearning.⁴

If one concentrates, however, on the dramatic changes within perception itself which, according to Kohler, occur during this adjustment, a theory of perception is needed which recognizes and deals with particular difficulties in interpreting the findings.

The principal difficulty concerns the fact that under identical retinal stimulation two or more different percepts may emerge, depending on different postural states of the organism. It does not appear possible to reconcile this finding with a sensorial perceptual theory; only a theory of perception which recognizes visual perception as an or-

ganismic rather than a sensory event can deal with this fact.

Once one has recognized that perception is an event depending on organismic factors (this would include not only postural, but emotive, motivational factors, etc.), the problem broadens to the more general question of how two essentially alien elements, such as muscular status of the organism on the one side and sensory processes on the other side, can interact with one another ("interaction paradox").

In sum, we submit (a) that Kohler's findings can be more adequately accounted for by an organismic than by a sensorial theory of perception; and (b) that an organismic theory adequate to interpret Kohler's findings has to contain constructs which are capable of handling the problem of the interaction of sensory and organismic factors.

SENSORY-TONIC FIELD THEORY OF PERCEPTION AND KOHLER'S FINDINGS

In this final section we should like to show how a particular organismic theory, which we have termed a sensory-tonic field theory of perception, interprets Kohler's findings. For this purpose, we will discuss first some of the pertinent aspects of sensory-tonic theory and then attempt to show how the essential findings of Kohler are consistent with the theory.

Sensory-tonic theory has several basic postulates employed in the explanation of perceptual events. The first postulate had its origin in the attempt to solve the so-called "interaction paradox," which pertains to the problem of how alien elements can interact. In order to solve this problem we propose that stimulation, whether it comes through extero-, proprio-, or interoceptors is sensory-tonic in nature. Thus, interaction between "muscular" and "visual" factors becomes understandable because both are fundamentally of

⁴ This sort of analysis leads to questions such as the existence of "postural" cues over and above "visual" cues, the relative importance of visual vs. postural cues, etc. It is not always realized that conclusions from experiments stated in terms of "cues" (e.g., "space perception is determined by postural as well as visual cues") are generalized descriptive statements about happenings without any commitments as to perceptual theory.

the same kind, viz., sensory-tonic in nature (21).⁵

A second postulate is that of "functional equivalence." This postulate asserts that diverse stimuli (visual, auditory, etc.), under ideal conditions, will lead to certain identical perceptual end products. We may refer here to a series of experiments which were designed to test the validity of the equivalence proposition. For instance, in one of our situations, the subject had the task of adjusting a rod in a darkroom so that it appeared vertical to him. The position at which a physical line was seen vertical is referred to as "apparent vertical." In these experiments the physical position of the apparent vertical differed under various conditions, e.g., under electrical stimulation of the right neck muscle it was tilted more to the left. Auditory stimulation functioned in a similar way, i.e., one-sided application shifted the position of apparent vertical opposite to the side to which the auditory stimulus was applied. Furthermore, related results were obtained under conditions of body tilt and rotation around the vertical axis of the body. When the body was tilted slightly to one side, the position of the apparent vertical, again, was tilted to the other side.⁶ Thus, these experi-

ments are in support of "functional equivalence" and therefore consistent with the first postulate of sensory-tonic field theory.

Another feature of sensory-tonic theory is its field-theoretical nature. The theory assumes that in dealing with perception we must consider the relationship between an object (psychophysical) and organism (psychophysiological). One will note that this view sets off sensory-tonic field theory not only from classical psychophysics, but also from orthodox gestalt psychology; as Brunswik has observed, orthodox gestalt psychology is "... as encapsulated within the organism as was ... classical psychophysics" (3, p. 63).

Thus, the third postulate is that a perceptual property is an experience which corresponds to a particular relation between organismic state and stimulus. We may re-emphasize that, in this view, perceptual experience does not correspond to an organismic state *per se* but to an organismic state related to impinging stimuli. Concretely: when stimuli impinge upon the organism, various basic relations are possible between stimuli and the ongoing momentary state of the organism. For example, one such relationship is that of stability, i.e., with a certain impinging stimulus there is no tendency for the pertinent aspects of the organismic state (sensory-tonic distribution) to change. Such a stable relationship between the impinging stimuli and the organismic state reflects itself in a particular percept. For instance, the perceptual experience is that of verticality if (ideally) there is symmetric sensory-tonic distribution in regard to a vertical axis of the erect body, and the pertinent stimuli are those coming from a perpendicular

⁵ We should like to point here to the evidence of sensory and tonic interdependence in perception by earlier investigators, notably Metzger (15), Goldstein (8), Schilder (17), Kleint (11), and Boernstein (1). For a brief review of this literature see Werner and Wapner (21). Within recent years one of the most lucid argumentations against a purely sensory theory of perception, from the viewpoint of neuropsychology, has been presented by Sperry (19). His notions on perceptual mechanisms, though clad in neurophysiological terms, have much in common with our own.

⁶ Prisms were used by Bruell (2) at the Clark laboratory as a technique of variation of oculo-muscular states. By this method Bruell was able to study interaction between

oculo-muscular states and retinal stimulation in regard to the position of the apparent median plane (straight ahead).

rod; in this situation, we assume that the sensory-tonic distribution in the organism does not tend to change, and it is such a stable organism-stimulus relationship which is mirrored in the experience of verticality.

In contrast, a relationship of instability occurs when certain stimuli issuing from an object instigate a tendency for the sensory-tonic distribution of the organism to change in a particular way. With regard to the dimension of verticality, this unstable physical relationship between the stimulus and organismic state reflects itself in the psychological experience of tilt.

In general, we may symbolize a stable relationship (R) by using the same subscript (x or y , etc.) for the pertinent stimulus property (s) and the pertinent aspect of organismic state (o). Thus, symbolically, a stable relationship is indicated by o_xRs_x , or o_yRs_y , etc. Correspondingly, an unstable relationship (R') may be symbolized as follows: $o_xR's_y$, or $o_yR's_x$, etc.

Coming back to one of our experiments mentioned above, the observer looks at a luminous rod in a dark room while the sternocleidomastoid neck muscle of the right side is stimulated. The stimulation of the neck muscle introduces a change in the organismic state from o_x (erect body position—without stimulation) to state o_y (erect body position—with stimulation to one side). A line in the physical position s_x , seen as vertical under state o_x , is no longer seen vertical in state o_y but is seen as tilted instead. There will be another physical position, s_y , which will provide a stable relationship in regard to o_y and will be perceived as vertical. Apparent verticality, previously a reflection of the relation o_xRs_x , has become now a reflection of the relation o_yRs_y .

A fourth postulate of sensory-tonic theory, critical here, refers to the relationship between what we have pre-

viously called an unstable relation ($o_xR's_y$) and a stable relation (o_yRs_y). The postulate states that given an unchanging stimulus (s_y) which is in an unstable relation to the existing organismic state (o_x), there is a tendency for the organism to change its state from o_x to o_y so that a higher degree of stability is obtained between stimulus and organism; that is, $o_xR's_y$ changes in the direction of o_yRs_y .

It seems to us that such facts as tendencies toward seeing slightly tilted lines as appearing progressively less tilted in the course of inspection (7, p. 453), and the progressive righting of the scene in a tilted mirror, etc., can be interpreted as examples of the stabilization tendency postulated above. That is, if the stimulus-organism relation under perception of tilt is that of $o_xR's_y$, then a tendency emerges for the organism to change its state from o_x to o_y so that o_yRs_y , which reflects the perceptual experience of verticality, ensues.

Having presented some basic assumptions and constructs of sensory-tonic field theory, we shall proceed to show how Kohler's findings can be translated rather easily into the conceptualizations of our theory.

Let us briefly review the sequence of events occurring in Kohler's experiments with half prisms. We may give a simplified description of the experiment by outlining five stages: Stage 1. This is the pre-experimental or control condition; here perception is undistorted, irrespective of whether the gaze is upwards or downwards. Stage 2. This represents a stage immediately following placement of the upper half prisms on the subject; here there is marked distortion when the subject looks through the prisms (gaze upwards), but relatively no distortion with downward gaze (no prism). Stage 3. At this stage, after having looked through the prisms for considerable time, there is relatively

TABLE 1
STIMULI, ORGANISMIC STATE, AND STIMULUS-
ORGANISM RELATION BETWEEN THEM FOR
THE FIVE STAGES AS VIEWED IN
SENSORY-TONIC TERMS

Direction of Gaze	Sequence of Events				
	1	2	3	4	5
	No Prism	Prisms On		Prisms Removed	
		Early	Late	Early	Late
Upwards	$o_x R s_x$	$o_y R' s_y$	$o_y R s_y$	$o_y R' s_x$	$o_x R s_x$
Downwards	$o_x R s_x$	$o_x R s_x$	$o_x R s_x$	$o_x R s_x$	$o_x R s_x$

little distortion either with upward (half prisms) or with downward (no prisms) gaze. Stage 4. At this stage immediately after removal of the half prisms there is a marked distortion in the opposite direction for gaze upwards (where prisms were present previously), and relatively little distortion with gaze downward. Stage 5. This stage refers to a phase some time later after removal of the prisms; now the condition of Stage 1 again prevails, i.e., there is distortion neither with gaze upwards nor with gaze downwards.

Table 1 presents, for each of the five stages, stimuli, organismic state, and stimulus-organism relation between them as viewed in sensory-tonic terms.

Let us first consider gaze upwards alone. In Stage 1, "normal" conditions are present; there is no perceptual distortion for gaze either upwards or downwards. In Stage 2, by virtue of the prism, the stimuli issuing from the object, have changed to s_y and the organism is still in state o_x . In Stage 3, after the subject has gazed through the prism for considerable time, s_y are still the stimuli and by virtue of adaptation, the organism has changed to state o_y . Stage 4 represents the relationships under initial removal of prisms; here the

removal of the prisms makes s_y become s_x and the organism is still in state o_y . Stage 5 represents the readaptation of the organism in terms of a change from o_y to o_x with stimuli s_x . Thus, with gaze upwards in Stages 2 and 4, the perception is distorted: $o_x R' s_y$, $o_y R' s_x$. In Stages 3 and 5, the perception is undistorted: $o_y R s_y$, $o_x R o_x$. Ideally, for gaze downwards, the relations throughout from Stage 1 through Stage 5 are $o_x R s_x$.

SUMMARY AND CONCLUSION

Up to now we have attempted to show the compatibility of Kohler's work with sensory-tonic field theory. But beyond this, the Innsbruck studies, though not designed within the framework of any particular theory, appear to us to advance the understanding of perception in sensory-tonic terms in three important ways:

1. In our own work, by manipulating the sensory-tonic state of observers we could demonstrate various effects of the organismic state upon perceived location and position of objects; on the other hand, though figural adaptation was clearly evidenced in our results, we could only infer rather than directly test the organismic factors involved in adaptation. Kohler's methodology bridges this gap: his device of halving the visual field into two areas, one distorted (gaze upwards), the other undistorted (gaze downwards), combined with the method of aftereffects, has led to results which demonstrated the involvement of organismic states (oculo-muscular states) in so-called visual adaptation. Under these experimental conditions, the ordinarily unified visual world became divided into two perceptual worlds depending on the oculo-muscular posture, i.e., the particular sensory-tonic states of the organism.

2. The role of organismic states was

tested in our experiments mainly in regard to such properties of objects as position and localization in space. In other words, these effects were mainly concerned with egocentric localization but not with "relative" position of objects in regard to one another. The Innsbruck studies have shown that figural adaptation and their aftereffects occur differentially with respect to direction of gaze; furthermore, they concern not only egocentric but also "relative" localization, as well as shape.⁷ Thus, these experiments give evidence in support of the role of sensory-tonic states in regard to a great variety of spatial properties.

Important as such evidence is for the generalization of organismic concepts, one word of caution should be added. If one chooses relatively simple situations, as we have done, certain rather specific hypotheses concerning underlying mechanisms can be made and tested; in contrast, the figural changes involved in Kohler's studies are of such complexity that a formulation of hypotheses on mechanisms specific to such changes would appear premature.

3. The third way in which Kohler's work significantly contributes to organismic theory hinges upon his work with color. His finding that color is affected analogously to spatial properties by organismic states broadens the conceptualization involved in sensory-tonic field theory to include a nonspatial property of visual perception. In this respect we should mention that we have already extended our experiments at least to one nonspatial property, viz., visual brightness. In several unpub-

lished studies⁸ organismic state was manipulated by means of "extraneous" tone stimulation; it was found that under these conditions brightness perception is affected. For instance, introduction of extraneous tone stimulation had an effect equivalent to the darkening of a grayish background against which a figure of a certain degree of brightness was perceived. The tentative conclusions we drew from these studies as to the sensory-tonic nature of brightness and its dependency on organismic states seem to find support and further generalization by the ingenious Innsbruck studies.

There will probably be those who would prefer to discuss Kohler's impressive findings in the familiar terms of "experience," "conditioned response," and the like. There need not be any disagreement about such interpretations if one considers them initial conceptualizations toward describing what happens behaviorally. We would, however, raise serious objections if this sort of interpretation should be called upon to replace any further analysis, and thus to obscure the important problem concerning the underlying perceptual mechanisms. That muscular or kinesthetic happenings, such as the position of the eye, can serve as a frame of reference within which color stimuli are experienced again brings up the "interaction paradox," viz., the problem of how seemingly alien factors, the muscular and the sensory, can interact with one another. This problem, it seems to us, has to be faced by any perceptual theory. The general conclusions from all of Kohler's experiments, especially from those on differential adaptation to color, seem to us strongly to support the central postulate of sensory-tonic theory that stimulation by any sensory factor is essentially a sensory-tonic

⁷ We may mention that one study conducted in the Clark laboratory concerned the perceptual property of shape (14); Kleint also has observed figural changes due to changes of tonus conditions; his results are, however, based mainly on introspective reports (11).

⁸ See in particular the theses by Mulholland (16) and by Gorrell (10).

(somato-tonic as well as viscerotonic) * event.

REFERENCES

1. BOERNSTEIN, W. On the functional relations of the sense organs to one another and to the organism as a whole. *J. gen. Psychol.*, 1936, 15, 117-131.
2. BRUELL, J. H. Visual egocentric localization: an experimental study. Unpublished doctor's dissertation, Clark University, 1953.
3. BRUNSWIK, E. The conceptual framework of psychology. *Int. Encycl. unified Sci.*, 1952, 1, 1-110.
4. ERISMANN, T. Das Werden der Wahrnehmung. *Tagungsbericht d. Berufsverbandes Deutscher Psychologen*, Bonn, 1947, 54-56.
5. EWERT, P. H. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior. *Genet. Psychol. Monogr.*, 1930, 7, 177-363.
6. GIBSON, J. J. Adaptation, after-effect and contrast in the perception of curved lines. *J. exp. Psychol.*, 1933, 16, 1-31.
7. GIBSON, J. J., & RADNER, M. Adaptation, after-effect and contrast in the perception of tilted lines. *J. exp. Psychol.*, 1937, 20, 453-467.
8. GOLDSTEIN, K. *The organism*. New York: American Book, 1939.
9. GOLDSTEIN, K., & ROSENTHAL, O. Zum Problem der Wirkung der Farben auf den Organismus. *Schweiz. Arch. f. Neurol. u. Psychiat.*, 1930, 26, 3-26.
10. GORRELL, R. B. The effect of extraneous auditory stimulation on critical flicker frequency. Unpublished doctor's dissertation, Clark University, 1953.
11. KLEINT, H. Versuche ueber die Wahrnehmung. *Z. Psychol.*, 1937, 140, 109-138; 142, 259-290; 1938, 143, 299-316.
12. KOHLER, I. Über Aufbau und Wandlungen der Wahrnehmungswelt. *Oesterr. Akad. d. Wissensch. Philos.-Histor. Kl.; Sitz.-Ber.*, 1951, 227, 1-118.
13. KOHLER, I. Umgewöhnung im Wahrnehmungsbereich. *Die Pyramide*, 1953, 5, 92-95; 6, 109-113.
14. KRUS, D. M. The effect of labyrinthian stimulation upon the perception of the shape of figures. Unpublished master's thesis, Clark University, 1951.
15. METZGER, E. Experimentelle Untersuchungen ueber den Lichttonus. *Arch. Ophthal. (Graefe)*, 1931, 127, 210-230.
16. MULHOLLAND, T. B. The effect of extraneous tone stimulation on visual brightness. Unpublished master's thesis, Clark University, 1953.
17. SCHILDER, P. *Brain and personality*. New York: International Universities Press, 1951.
18. SNYDER, F. W., & PRONKO, N. H. *Vision with spatial inversion*. Wichita: University of Wichita Press, 1952.
19. SPERRY, R. W. Neurology and the mind-brain problem. *Amer. Scientist*, 1952, 40, 291-312.
20. STRATTON, G. M. Vision without inversion of the retinal image. *Psychol. Rev.*, 1897, 4, 341-360; 463-481.
21. WERNER, H., & WAPNER, S. Sensory-tonic field theory of perception. *J. Pers.*, 1949, 18, 88-107.
22. WERNER, H., & WAPNER, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, 59, 324-338.
23. WERNER, H., WAPNER, S., & BRUELL, J. H. Experiments on sensory-tonic field theory of perception. VI. Effect of position of head, eyes and of object on position of the apparent median plane. *J. exp. Psychol.*, 1953, 46, 293-299.

* With respect to the viscerotonic involvement in visual perception, Goldstein's observations (9) on specific organic responses to "cold" and "warm" color are highly suggestive.

(Received March 22, 1954)

SINCE LEARNED BEHAVIOR IS INNATE, AND VICE VERSA, WHAT NOW?¹

WILLIAM S. VERPLANCK

Harvard University

My text for today may be found in the behavioral apocrypha; it pertains to the courtship of the three-spined stickleback. When the male stickleback, with nest completed, and in full bloom of red belly and bright blue eye, is confronted with a silvery, deep-bosomed, egg-laden female, he courts her. He zigs toward her, zags back, and toward and back again, approaching nearer and nearer with successive passes. After his closest approach, he swims directly toward the entrance to his nest. If she fails to follow, he turns away and repeats his game—and again and again, until she follows. And when she acquiesces, when she does follow, his behavior changes; this time, he reaches the nest entrance, he swims into it and straight through the nest, leading the female.

One studies this behavior by presenting the male with a captive female, one that cannot follow because the experimenter has confined her in a glass test tube. After a series of presentations of the female in this way, with repeated elicitations of the zigzag dance of the male, it is possible to perform a kind of control experiment in which the male is presented with an empty test tube. As a control, this usually proves disappointing; the male, as like as not, will proceed to zigzag, to court the test-tube. But what is more interesting, I am told that

if the experimenter is compliant, and moves the test tube like a female in pursuit of the male as he swims directly toward the nest, the male will proceed with his usual behavior toward the female; he will swim into the nest, and through it.

This is one of the kinds of things the ethologists² bring to us. With their background in biology, and their strictly behavioristic approach, they have over the past few years staggered us with a mass of data (10, 11; see also *Behaviour* and the *British Journal of Animal Behaviour*) on the behavior of inframammals, and have provoked once again discussion of an issue that from time to time drags psychologists out of the laboratory and into the forum—nature or nurture, heredity or environment, learning or maturation.

Ethologists, of course, stress “innate” behavior; they find differences between the behaviors they study, and those learning-theory psychologists work on. These they attribute to innateness. That such innateness as a source of differences is hopelessly confounded with differences in the species studied, and with

¹ A paper presented at the symposium titled “Unlearned Behavior: Concepts and Findings,” sponsored jointly by Divisions 3 and 7 of the American Psychological Association, in September, 1954. It was made possible by a grant from the American Philosophical Society, which enabled the writer to gain first-hand familiarity with the work of ethologists.

² Ethology, as it is defined by ethologists, is the science of behavior. And so it is: the methods and concepts of ethology bear a very close resemblance to their counterparts in the modern behaviorism psychologists are familiar with. This is true despite a difference in subject matter; ethologists, trained in the intellectual tradition of zoology, have concerned themselves largely with species-specific behavior in birds and fishes, and the individual (“acquired”) behavior of insects (especially wasps). One ethologist explained ethology this way: “Ethologists are behaviorists who like their animals.”

the degree of experimental restriction on the behavior investigated, goes without saying.

Ethologists have, however, stirred up the argument, and one of the features of our cultural heritage appears to be a very strong tendency to take one or another side in it. The discussion shifts its ground, changes its dominant features, but it seems to remain.

Fortunately, there are signs that progress is being made toward a solution to the endless chatter. Experimental animals, busily engaged in coping with their environments, seem to be coping with this bit of human activity as well. They seem to be telling us that there is no possible reasonable answer to questions such as "Is this bit of behavior innate?" The dilemma of innate or acquired seems to be one of those categorical pseudo problems that the philosopher Ryle (9) has concerned himself with. Analysis of the variables controlling behavior, such as that presented by Hebb (4), tells us the same thing: it is a logical impossibility to study the innate before studying learning.

As we list the criteria proposed for distinguishing the innate from the learned, many examples of learned behavior can be found that satisfy one or more—or most—of the criteria for the innate. Stereotypy, universality of appearance, orderliness, adaptivity, resistance to modification—all these fail, and only one criterion remains: execution of the behavior on its first opportunity to occur, *without* the possibility of previous learning. *How* does one rule this out?

We are forced into the position of acknowledging that the only criterion for distinguishing between innate and acquired behavior is one that requires us to accept the null hypothesis as proven.

We are forced into the position that no meaningful distinction can be drawn

between learned and innate behavior, that is, between the stereotyped and highly predictable behavior studied in inframammals by the ethologists, and the more variable behavior studied in the T maze and Skinner box. We can no more distinguish between behavior that is innate and behavior that is learned than physicists can distinguish between light that is made up of corpuscles and light that is made up of transverse vibrations. Innate and learned must always be assumed to have quotation marks about them.

Given this position—the assumption that much the same behavioral laws apply throughout the vertebrate realm (if not further), together with a rigorous experimental approach to behavior—a number of consequences follow. I shall note and illustrate some of these consequences.

First, the same classes of experimental variables should control both learned and unlearned behavior,³ and in similar ways. Hence, the same kinds of difficulties should arise in experimental designs aimed at untangling the effects of these variables, whether they are applied to learned or innate behavior.

Secondly, the structure of the theory of unlearned behavior and that of learned behavior must prove to be similar if not identical. It should not matter which kind of behavior yields a theory—it should apply to both. That is to say, one theoretical account of causal variables should be found forced by the data, and it should prove applicable to the data of both. And perhaps the same classes of theoretical variables should give the same kind of theoretical difficulties.

Thirdly, the behavioral phenomena

³ From this point on, by learned behavior, we will refer to behavior studied by learning theory men; and by unlearned or innate behavior, to behavior studied by ethologists, and so termed by these men.

observed by ethologists in the field should be observable in the T maze and similar apparatus, and those observed in the T maze should be observable in the field. Those who study innate behavior are well advised to look for and control the effects of the variables studied in learning. Conversely, those who study learning may facilitate the accumulation of relevant data by looking for and controlling the effects of the variables that have been considered the preserve of the ethologist. I am not saying that ethologists and psychologists should borrow each other's concepts to apply to their own fields, but rather that there is *only one* field of investigation, and failure to examine all the concepts developed in that field may serve to delay the development of an ordered and comprehensive body of data.

If there were time, we could develop fully each of these points. We could draw the revealing parallels between the series of experiments on generalization gradients in the CR (e.g., 5, 6) and those on the sign stimuli releasing pecking in the gull chick (12). The differences seem to lie in the number of dimensions of the stimulus that must be investigated. Hovland restricts himself to one: metric distance from the point of original conditioning along the body surface. Others have explored response strength along other single stimulus dimensions, such as frequency of a pure tone. Tinbergen and Perdeck, however, are forced by the animal, as it were, to vary the complex stimulus for the pecking response along a series of dimensions that includes those relating to the size, position, and optical properties of the spot on the bill of an adult sea gull, as well as to similar measures of the bill itself and of the remainder of the adult's head. Along each of these dimensions they find what is recognizable as a generalization gradient. To determine these gradients, careful counterbalancing of

the order of test stimuli was essential, lest, as in the CR, the response extinguish.

We could point out that it does no abuse to the theories now current in ethology to change a term here and there and thereby convert ethological theory into something in large measure almost indistinguishable from 1943 Hullian theory (7), except in its treatment of the variables defining sH_R (some of which it ignores). We could expand the difficulties that ethological and learning theorists have with the concepts of stimulus and response. Both tend to commit a series of equivocations or, more charitably, both tend to gloss over some of the possible distinctions to be drawn between movement-defined and consequence-defined responses (3, 11). More profitably, we may look at some of the contributions that the work of the ethologists may make to behavioral psychology today.

After Beach's classic paper, "The Snark Was a Boojum" (2), it may seem a bit unnecessary to stress once again that behavior is a function of species membership, and that we do not have as yet more than the beginnings of a comparative psychology. The neglect of this field has served to let the work of psychologists drift out of the context of that general evolutionary theory that unifies all the biological sciences. We have little to say about the survival value⁴ and the taxonomic distribution of behavior. This may or may not be a bad state of affairs. But clearly, in forgetting the comparative approach, we are very sharply limiting the inductive basis of our theories of behavior. To restrict ourselves to the study of one or two species and a short list of en-

⁴ Most psychologists are surprised when they learn that the adaptiveness or survival value of a structure or of a behavior is now subject to experimental verification in biology.

vironmental variables ensures that miniature theories will remain miniature.

The penalty we pay can be evaluated very readily indeed when once we shift from species to species. Armed with our knowledge of rats, we may attempt, for example, to do an experiment on the guinea pig. As soon as we do this, we discover that guinea pigs are not rats. This is not very startling, perhaps, but significant enough when we discover that the food-deprivation-feeding-behavior relationships⁵ in the guinea pig are sufficiently different from those in the rat that it is all but impossible to recover from the guinea pig the types of relationships between bar pressing and deprivation that we are familiar with in the rat. Here is one area of comparative psychology that has hardly been touched, but one that has many implications, both practical and theoretical. The laws of learning operate within limits determined by the genetic characteristics of the animal. The behavior repertoire and the functional relationships existing among members of the repertoire and between them and the environment are set by the genetic characteristics of the animal, whether they are learned or not.

Secondly, we might look toward incorporating into our theories of learning a variable that has been remarkably neglected—in theories if not in practice. This variable is the *age* of the organism, and the general topic is "the critical period for learning." But with D. O. Hebb and Eckhard Hess among the participants in this symposium, it seems hardly necessary to stress the fact that the acquisition of particular modes of behavior is very much a function of the animal's age, of behavior acquired before the behavior we are studying is

acquired, and of other behavior being acquired, or appearing, at the same time.

Thirdly, psychologists have largely set problems to their animals, and examined the incidence of particular pieces of behavior. Starting out with limited interests, we have examined limited sets of variables, both dependent and independent. Ethologists, on the contrary, have often examined somewhat broader causal chains of events, and have not restricted themselves to observing the effects of *our* sets of independent variables on the responses *we* like to deal with. To put it another way, ethologists watch everything that animals do, and everything that is happening when they do it. Given the opportunity, animals will emit behavior that is highly adaptive, and integrate it into longer causal chains, still under the control of specific motivating conditions and contingencies of reinforcement. Quite interesting behavioral effects will occur if we but let them. Let me take two examples.

In the first case, worked on by O. R. Lindsley, 20 rats were put on a deprivation schedule that allowed them access to solid Purina chow pellets for just one hour a day. Lindsley placed a number of pellets in the individual cages at the beginning of the hour, and at the end of the hour he took the remaining ones out. The first day was uneventful. There was the rat, there were the pellets, and in went the hand to remove the pellets. On the second day, when the experimenter's hand moved into the cage of *one* animal, the rat seized a pellet and retreated to the rear of the cage with it. There he held it. He hung on. After taking out the other pellets, the experimenter had to take the pellet from the rat forcibly. On the next day, two or three rats showed this behavior, and by the end of the week, not only were all 20 showing *at least* this behavior, but some had gone on to seize and carry two or more pellets to the rear, and in

⁵ This and several other experimental results described have not yet been reported or published elsewhere.

many the behavior was set off by simply unlatching the cage door. Reinforced? Interesting? Irrelevant? Some data we now have suggest that this behavior is associated with an increased tendency to run in a running wheel.

In the second case, 15 rats were trained to bar press in simple wire mesh Skinner boxes. After they were placed on a reinforcement schedule, discrimination training was begun, with five-minute periods of presentation of S_D , the positive discriminandum, and five minutes of S_A , the negative discriminandum. S_A was a two-inch equilateral black cardboard triangle that could be fixed by paper clips to the wire mesh behind the bar of the Skinner box. Under S_D conditions, this black triangle was absent. During the first discrimination period with the first failures of reinforcement, most of the animals seized, shook, and bit the bar, and the black triangle too. In most cases this aggressive behavior continued for several minutes; it continued in fact until the triangle was removed from behind the bar, usually by the rat. As soon as the triangle was gone, the animals resumed bar pressing at their normal rates. At the end of a series of such cycles of S_D and S_A , the majority of the rats, as soon as S_A was fixed in place, would approach it, and neatly but vigorously remove it—to get back to work. The effect appeared whether or not reinforcements were forthcoming when S_A was removed, although it is not surprising that the behavior was not so neat and efficient when the recurrence of S_D did not restore the previous reinforcement schedule. Altogether there was a startling shift from "violent," "aggressive," "emotional" assault to neat, smooth, precise manipulation.

Both these cases are the kinds of behavior that ethologists look for, and with which they deal. Both these cases are the kind of behavior that psychologists

learn to ignore, or learn to prevent. Perhaps I may be forgiven if I suggest that they should be of some concern and interest to the learning theorist.

The next area that might be examined is that of displacement activities—responses usually associated with one drive state that appear when behavior associated with another drive is prevented from occurring, and in conflict situations (1). One mode of behavior that might be classified as displacement activity by the ethologist is the bar biting and bar shaking of our last example. Others are the face washing and general body cleaning—preening—that rats so often vigorously show during extinction, and the cooing and wing beating of the pigeon that occur under the same circumstances. These modes of behavior are familiar enough, but they are very largely associated for us with the extinction of behavior that has led to food and water. It is clear that our animals have behaviors that are not associated with reinforcement by food and water getting, or by shock avoidance. It has now been shown that mice will learn to press bars when the reinforcement is access to a submissive male (which is then attacked), that rats will learn to press bars that release brakes on their running wheels (8) or that permit a female in oestrus to join them in the Skinner box. What behaviors, if any, appear when extinction is carried out with respect to such reinforcements? The ethological work would suggest strongly that we look and see. Perhaps some things that have puzzled us will be clarified.

Again, much work has been done with the hoarding of food and water pellets, and of other objects, under experimental conditions involving deprivation of these substances. But what of rats that hoard yellow pencils, pens, paper clips, and envelopes, when they have not had a history of deprivation of food or water, and when such objects have not been

associated with food or water? And what of satiated rats running freely about familiar spaces that transport large pieces of favored foods to preferred places, and there eat them, although they eat small pieces on the spot when they get them? And that do the same with their standard food if they are hungry? And that do not hoard it? The psychologist's fixation on a few easily manipulated and already identified experimental variables may seriously limit the power of his generalizations.

Let me complete the circle, and close by returning to the behavior described in the beginning, that is, to the stickleback that was conditioned to zigzag to a test tube. When this conditioning was effective, he showed other behavior toward the test tube—behavior that had not been conditioned to it. He led it, you will recall, to his nest, and there he did what he usually does only when a female is behind him: he swam through the nest. I cannot think of a clear-cut case where such a transfer effect occurs in the animals run by psychologists. What is disturbing is that I cannot think of a clear-cut case where such a transfer effect *could* have been observed in psychologists' animals.

Some rather similar behaviors occur to me, and these are ones that have puzzled me as an S-R psychologist. Let us take rats that have run across a grid to food many times, so that their running is asymptotic. If, for the first time, they are now shocked as they cross the grid, they may show no change in their running when the shock occurs. But on the *next* trial, they may stop short of the grid, and there sniff and poke and peer.

The ethologists have much to teach us. . . .

They suggest that we must *observe*

more species, that we must find new organizations of behavior (that is, new drives), that we must permit more manipulation of the environment by the animal, giving opportunities for phenomena not already in the textbooks to occur—in sum, that we must broaden the inductive basis, the experimental underpinnings, of our theories of behavior.

REFERENCES

1. BASTOCK, M., MORRIS, D., & MOYNIHAN, M. Some comments on conflict and thwarting in animals. *Behaviour*, 1953, 6, 66-84.
2. BEACH, F. A. The snark was a boojum. *Amer. Psychologist*, 1950, 5, 115-124.
3. ESTES, W. K., et al. *Modern learning theory*. New York: Appleton-Century-Crofts, 1954.
4. HEBB, D. O. Heredity and environment. *Brit. J. Anim. Behav.*, 1953, 1, 43-47.
5. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, 17, 125-148.
6. HOVLAND, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, 51, 279-291.
7. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
8. KAGAN, J., & BERKUN, M. The reward value of running activity. *J. comp. physiol. Psychol.*, 1954, 47, 108.
9. RYLE, G. *Dilemmas*. Cambridge: Cambridge Univ. Press, 1954.
10. SOCIETY FOR EXPERIMENTAL BIOLOGY. *Physiological mechanisms in animal behaviour*. New York: Academic Press, 1950.
11. TINBERGEN, N. *The study of instinct*. Oxford: Oxford Univ. Press, 1951.
12. TINBERGEN, N., & PERDECK, A. C. On the stimulus situation releasing the begging response in the newly hatched Herring Gull chick (*Larus argentatus argentatus* Pont.). *Behaviour*, 1950, 3, 1-39.

(Received for early publication October 20, 1954)

ZEITSCHRIFT FÜR EXPERIMENTELLE UND ANGEWANDTE PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

This journal is devoted to experimental investigations covering the entire field of psychology. In addition, up-to-date reviews of current books on psychology are published.

For many years, a strictly scientific journal of psychology has been lacking in German-speaking countries. In taking over this task the "Zeitschrift für experimentelle und angewandte Psychologie" endeavours to continue the tradition of the former, and now extinct, German journals of scientific psychology.

All articles are provided with summaries in English and French.

Herausgeber für Deutschland

Prof. Dr. J. von Allesch, Göttingen; Prof. Dr. Ph. Lersch, München

Herausgeber für Österreich

Prof. Dr. H. Rohracher, Wien

Herausgeber für die Schweiz

Prof. Dr. R. Meili, Bern

Redaktion

Dr. C. J. Hogrefe, Göttingen

Ständige Mitarbeiter

Prof. Dr. W. Arnold, Nürnberg; Prof. Dr. H. Biäsch, Zürich; Prof. Dr. C. Bondy, Hamburg; Prof. Dr. E. Brunswik, Berkeley USA; Prof. Dr. H. Düker, Marburg; Prof. Dr. G. Ekman, Stockholm; Prof. Dr. Th. Erismann, Innsbruck; Prof. Dr. H. J. Eysenck, London; Prof. Dr. R. Heiß, Freiburg i.B.; Prof. Dr. B. Herwig, Braunschweig; Prof. Dr. H. Hetzer, Marburg; Prof. Dr. E. Kretschmer, Tübingen; Prof. Dr. O. Kroh, Berlin; Prof. Dr. W. Metzger, Münster; Prof. Dr. P. Pichot, Paris; Prof. Dr. E. Rausch, Frankfurt; Prof. Dr. G. Révész, Amsterdam; Prof. Dr. E. Sander, Berlin; Prof. Dr. U. Undeutsch, Köln; Prof. Dr. A. Wellek, Mainz; Prof. Dr. K. Wilde, Göttingen.

The journal appears quarterly and contains 160 pages.

Subscription is DM 64.— per volume and DM 18.— for single numbers.

VERLAG FÜR PSYCHOLOGIE • DR. C. J. HOGREFE • GÖTTINGEN
(Postfach 414, Germany)

The first comprehensive, systematic survey

THEORIES OF PERCEPTION AND THE CONCEPT OF STRUCTURE

**A Review and Critical Analysis with an Introduction
to a Dynamic-Structural Theory of Behavior**

By FLOYD H. ALLPORT, *Syracuse University*

This unique, double-purpose volume provides the first comprehensive, organized review of the major current theories of perception, and offers a foundation for future systematic work in the form of a general theory of structure in behavior.

The book opens with a review of the essential facts and theories of perception, critically examining the theory in the light of scientific method. It then covers each of the major theories separately and systematically, and offers a synthesis made up of the 8 generalizations more or less common to all. The need for an over-all explicit theory of structure is demonstrated. Such a theory is tentatively outlined in the concluding chapter. 1955. 709 pages. \$8.00.

Send now for a copy on approval

JOHN WILEY & SONS, Inc. 440 Fourth Ave., New York 16, N.Y.

WRITING A BOOK?

Then you too must have pondered the question, as so many new authors have: "How can a significant work which is not necessarily a guaranteed commercial success, or a candidate for the best-seller list, be published?"

Our extensive experience in regular commercial and subsidy publishing has made clear to us the need for a 100% honest, selective, and professionally skilled cooperative publisher. **THIS IS THE FUNCTION WE FULFILL.** Our books are handsomely designed, carefully edited, and intelligently promoted. Our books sell, not in explosive spurts, but steadily and regularly. Our imprint is esteemed by libraries, bookstores, reviewers, and literary agents.

Many titles are published on some form of cooperative basis with higher royalties. Send your manuscript, without obligation, for editorial evaluation.

THE AMERICAN PRESS, INC.

Atten: Mr. Revere

489 Fifth Ave.

New York 17, N. Y.